ZXR10 8900 Series
10 Gigabit Routing Switch
User Manual (MPLS Volume)

Version 2.8.02.C

ZTE CORPORATION
NO. 55, Hi-tech Road South, ShenZhen, P.R.China
Postcode: 518057
Tel: (86) 755 26771900
Fax: (86) 755 26770801
URL: http://ensupport.zte.com.cn
E-mail: support@zte.com.cn

LEGAL INFORMATION

Copyright © 2010 ZTE CORPORATION.

The contents of this document are protected by copyright laws and international treaties. Any reproduction or distribution of this document or any portion of this document, in any form by any means, without the prior written consent of ZTE CORPORATION is prohibited. Additionally, the contents of this document are protected by contractual confidentiality obligations.

All company, brand and product names are trade or service marks, or registered trade or service marks, of ZTE CORPORATION or of their respective owners.

This document is provided "as is", and all express, implied, or statutory warranties, representations or conditions are disclaimed, including without limitation any implied warranty of merchantability, fitness for a particular purpose, title or non-infringement. ZTE CORPORATION and its licensors shall not be liable for damages resulting from the use of or reliance on the information contained herein.

ZTE CORPORATION or its licensors may have current or pending intellectual property rights or applications covering the subject matter of this document. Except as expressly provided in any written license between ZTE CORPORATION and its licensee, the user of this document shall not acquire any license to the subject matter herein.

ZTE CORPORATION reserves the right to upgrade or make technical change to this product without further notice.

Users may visit ZTE technical support website http://ensupport.zte.com.cn to inquire related information.

The ultimate right to interpret this product resides in ZTE CORPORATION.

Revision History

Revision No.	Revision Date	Revision Reason
R1.3	20100205	Edition update

Serial Number: sjzl20093840

Contents

About This Manual I
Safety Instruction 1
Safety Instruction 1
Safety Signs 1
MPLS Configuration 3
MPLS Overview 3
MPLS Function 3
IP Forwarding Features 4
ATM Forwarding Features 4
MPLS Features 5
MPLS Working Principle 6
MPLS Terms 7
MPLS Label 7
Forward Equivalence Class 8
MPLS Network Structure 9
LSP Establishment 9
LSP Establishment Procedure 10
Penultimate Hop Popping 12
Label Control and Management 13
Label Distribution Mode 14
Label Control Mode 15
Label Holding Mode 15
Label Forwarding Table 16
Label Distribution Protocol 16
LDP Session Establishment and Maintenance 17
Configuring MPLS 18
Enabling MPLS 18
Configuring LDP Router-ID 18
Configuring Label Distribution and Control 19
Configuring LDP Time Parameters 20
Configuring MD5 Password 20

Clearing LSR Peer Connection....20
MPLS Configuration Examples....21
MPLS Basic Configuration Example....21
Label Distribution Example for Specified Nerwork Segment....23
MPLS Maintenance and Diagnosis....25
MPLS L3VPN Configuration....29
MPLS L3VPN Overview....29
Related Terms....30
VPN-IPv4 Address and Route Distinguisher....31
MPLS L3VPN Working Principles....32
MPLS L3VPN Load Balance....33
MPLS L3VPN Load Balance on Smartgroup Link....34
Configuring MPLS L3VPN....35
Configuring OSPF in MPLS Domain....35
Configuring VRF on PE....36
Configuring Static Route between CE and PE....38
Configuring RIP between CE and PE....39
Configuring OSPF between CE and PE....40
Configuring EBGP between CE and PE....42
Configuring MPBGP....43
Configuring MPLS L3VPN Load Balance....45
MPLS L3VPN Configuration Example....45
Configuration Example of MPLS L3VPN Load Balance on Smartgroup Link....49
MPLS L2VPN Configuration....51
MPLS L2VPN Overview....51
Virtual Private Wire Service....51
Virtual Private LAN Service....52
Configuring MPLS L2VPN....54
Configuring VPWS....54
Configuring VPLS....54
Configuring Layer 2 Control Message Transparent Transmission VPLS Network Policy....56
MPLS L2VPN Configuration Examples....57
VPWS Configuration Example....57
VPLS Configuration Example....58
Layer 2 Control Message Transparent Transmission VPLS Network Configuration Example....62
MPLS L2VPN Maintenance and Diagnosis....62

VPWS Maintenance and Diagnosis 62
VPLS Maintenance and Diagnosis 63
Layer 2 Control Message Transparent Transmission VPLS Network Configuration Maintenance and Diagnosis 63
MPLS OAM Configuration 65
Function Introduction 65
Configure MPLS OAM 66
MPLS TE HOT-STANDBY Configuration 69
MPLS TE HOT-STANDBY Overview 69
Basic Configuration of MPLS TE HOT-STANGBY 69
Maintenance and Diagnosis of MPLS TE HOT-STANDBY 70
Configuration Example of MPLS TE HOT-STANDBY 70
Figures 73
Glossary 75

This page is intentionally blank.

About This Manual

Purpose

This manual is ZXR10 8900 Series (V2.8.02.C) 10 Gigabit Routing Switch User Manual (MPLS Volume) and applies to ZXR10 8902/8905/8908/8912 10 Gigabit routing switch.

Intended Audience

This manual is intended for the following engineers:

- On-site maintenance engineers
- Network monitoring engineers
- System maintenance engineers

What Is in This Manual

This manual contains the following contents:

Chapter	Overview
Chapter 1 Safety Instruction	Describes safety instructions and signs used in this manual.
Chapter 2 MPLS Configuration	This chapter describes basic concept, configuration and configuration example of MPLS.
Chapter 3 MPLS L3VPN Configuration	This chapter describes basic concept, configuration and configuration example of MPLS L3VPN.
Chapter 4 MPLS L2VPN Configuration	This chapter describes basic concept, configuration and configuration example of MPLS L2VPN.
Chapter 5 MPLS OAM Configuration	This chapter describes basic concept, configuration and configuration example of MPLS OAM.
Chapter 6 MPLS TE HOT-STANDBY Configuration	This chapter describes basic concept, configuration and configuration example of MPLS TE HOT-STANDBY.

Related Documentation

The following documentation is related to this manual:

- ZXR10 8900 Series (V2.8.02.C) 10 Gigabit Routing Switch Hardware Installation Manual
- ZXR10 8900 Series (V2.8.02.C) 10 Gigabit Routing Switch Hardware Manual
- ZXR10 8900 Series (V2.8.02.C) 10 Gigabit Routing Switch User Manual (Basic Configuration)
- ZXR10 8900 Series (V2.8.02.C) 10 Gigabit Routing Switch User Manual (Ethernet Switching)
- ZXR10 8900 Series (V2.8.02.C) 10 Gigabit Routing Switch User Manual (IPv4 Routing)
- ZXR10 8900 Series (V2.8.02.C) 10 Gigabit Routing Switch User Manual (IPv6)

- ZXR10 8900 Series (V2.8.02.C) 10 Gigabit Routing Switch User Manual (MPLS)
- ZXR10 8900 Series (V2.8.02.C) 10 Gigabit Routing Switch User Manual (DPI)
- ZXR10 8900 Series (V2.8.02.C) 10 Gigabit Routing Switch User Manual (Firewall)
- ZXR10 8900 Series (V2.8.02.C) 10 Gigabit Routing Switch Command Reference (Index)
- ZXR10 8900 Series (V2.8.02.C) 10 Gigabit Routing Switch Command Reference (Basic Configuration Volume I)
- ZXR10 8900 Series (V2.8.02.C) 10 Gigabit Routing Switch Command Reference (Basic Configuration Volume II)
- ZXR10 8900 Series (V2.8.02.C) 10 Gigabit Routing Switch Command Reference (Basic Configuration Volume III)
- ZXR10 8900 Series (V2.8.02.C) 10 Gigabit Routing Switch Command Reference (Ethernet Switching)
- ZXR10 8900 Series (V2.8.02.C) 10 Gigabit Routing Switch Command Reference (IPv4 Routing Volume I)
- ZXR10 8900 Series (V2.8.02.C) 10 Gigabit Routing Switch Command Reference (IPv4 Routing Volume II)
- ZXR10 8900 Series (V2.8.02.C) 10 Gigabit Routing Switch Command Reference (IPv6)
- ZXR10 8900 Series (V2.8.02.C) 10 Gigabit Routing Switch Command Reference (MPLS)
- ZXR10 8900 Series (V2.8.02.C) 10 Gigabit Routing Switch Command Reference (Network Management)
- ZXR10 8900 Series (V2.8.02.C) 10 Gigabit Routing Switch Command Reference (QoS)
- ZXR10 8900 Series (V2.8.02.C) 10 Gigabit Routing Switch Command Reference (Security)
- ZXR10 8900 Series (V2.8.02.C) 10 Gigabit Routing Switch Command Reference (Voice and Video)
- ZXR10 8900 Series (V2.8.02.C) 10 Gigabit Routing Switch Command Reference (Multicast)
- ZXR10 8900 Series (V2.8.02.C) 10 Gigabit Routing Switch Command Reference (DPI)
- ZXR10 8900 Series (V2.8.02.C) 10 Gigabit Routing Switch Command Reference (Firewall)

Chapter 1

Safety Instruction

Table of Contents

Safety Instruction 1
Safety Signs 1

Safety Instruction

Only duly trained and qualified personnel can install, operate and maintain the devices.

During the device installation, operation and maintenance, please abide by the local safety specifications and related operation instructions, otherwise physical injury may occur or devices may be broken. The safety precautions mentioned in this manual are only supplement of local safety specifications.

The **debug** commands on the devices will affect the performance of the devices, which may bring serious consequences. So take care to use **debug** commands. Especially, the **debug all** command will open all **debug** processes, so this command must not be used on the devices with services. It is not recommended to use the **debug** commands when the user networks are in normal state.

ZTE Corporation will assume no responsibility for consequences resulting from violation of general specifications for safety operations or of safety rules for design, production and use of the devices.

Safety Signs

The contents that users should pay attention to when they install, operate and maintain devices are explained in the following formats:

Warning:

Indicates the matters needing close attention. If this is ignored, serious injury accidents may happen or devices may be damaged.

Caution:

Indicates the matters needing attention during configuration.

Note:

Indicates the description, hint, tip, and so on for configuration operations.

Chapter 2

MPLS Configuration

Table of Contents

MPLS Overview 3
MPLS Terms 7
LSP Establishment 9
Label Control and Management 13
Configuring MPLS 18
MPLS Configuration Examples 21
MPLS Maintenance and Diagnosis 25

MPLS Overview

MPLS Function

With the rapid development of network scale and user number, people pay more attention to the problem that how to extend service further on the network and how to improve the quality of services. As the IP protocol is a connectionless protocol, it is hard to provide enough throughputs and time delay according to the requirements. For IP protocol, it just uses best-effort mechanism to satisfy the users. Without other measures to improve the current network conditions, it is hard to develop new services greatly.

Among existing network technologies, ATM has advantages for quality of services. However, the high cost of ATM network makes it hard to implement. In addition, although ATM switches are widely used as backbone nodes in networks, services that take ATM cells to desktop develop very slowly.

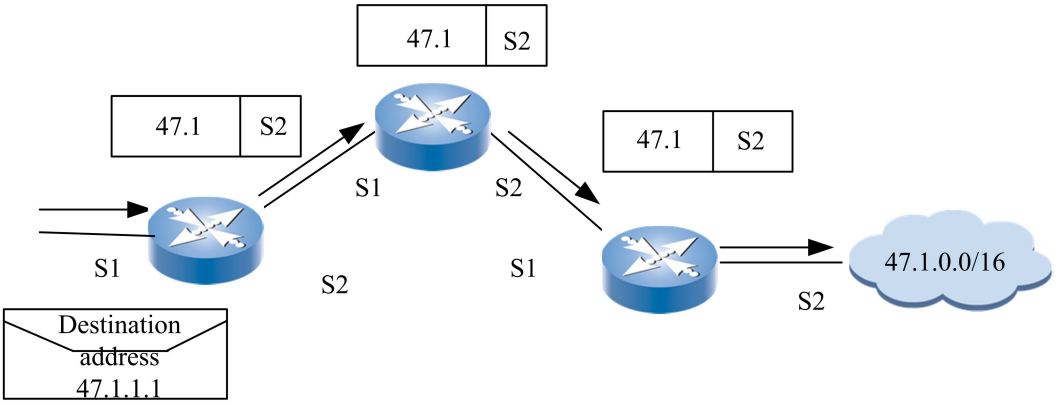
IP and ATM technologies meet difficulties with respective field, and they need each other to have further development. There is inevitability to combine the two technologies. Therefore, Multi-Protocol Label Switching (MPLS) technology comes into being. It integrates the advantages of switching technology in network core and IP routing technology in network edge.

MPLS uses label switching, so routers only need to judge the labels when they forward packets. MPLS supports protocols in network layer (such as IPv6, IPX and IP) as well as protocols in data link layer (such as ATM, FR and PPP).

IP Forwarding Features

Features of traditional IP forwarding is shown in Figure 1.

FIGURE 1 IP FORWARDING



IP uses hop-by-hop method for communication. When a packet reaches a router, the router looks up its routing table to forward the packet to the next router. After forwarding, the packet is not controlled by the router any longer. When there is correct routing information in the routers along the whole path, the packets can be forwarded correctly.

When looking up the routing table, the router uses the longest-matching principle. Therefore, it is difficult to implement fast forwarding.

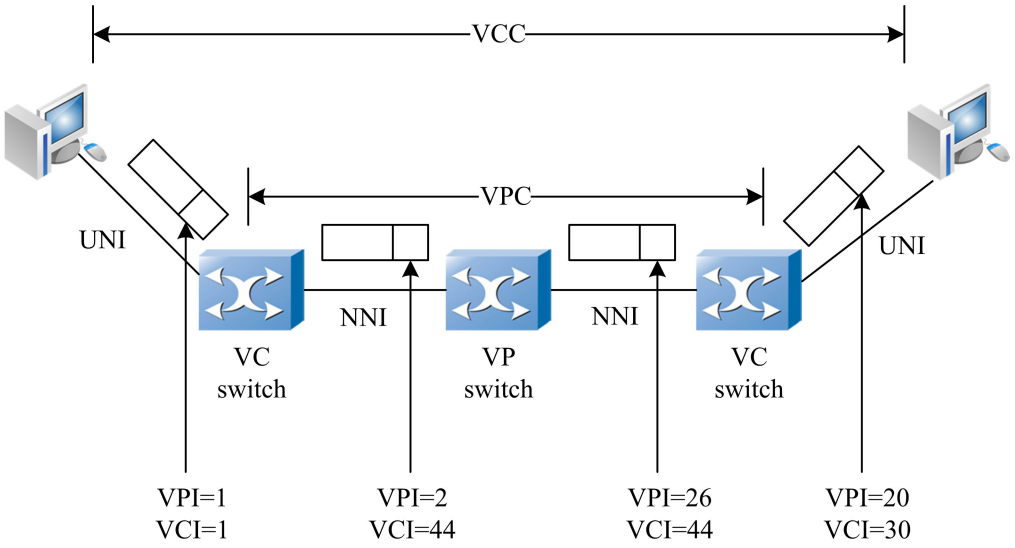
Routers have to collect information of different network segments for different routing protocols. Otherwise, the router can not forward the packets.

IP network is a connectionless network, therefore its QoS is not well ensured.

ATM Forwarding Features

ATM forwarding procedure is shown in Figure 2.

FIGURE 2 ATM FORWARDING



ATM switches forward packets through VPI/VCI. The switches do not judge routing information when forwarding ATM cells.

When an ATM switch receives a packet, it looks up VPI/VCI table and forwards the packets in switching way.

When forwarding packets, ATM switches have the following features:

- ATM chooses the path on the base of data link layer. VPI/VCI is effective for local switch. The fast looking-up is implemented by hardware.
- ATM network is a connection-oriented network, so it can implement QoS according to different VPIs/VCIs.
- ATM supports traffic control mechanism and different kinds of services, such as real time service.

MPLS Features

MPLS belongs to the third generation of network architecture. It is a new switching standard of IP high-speed backbone network, brought forward by Internet Engineering Task Force (IETF).

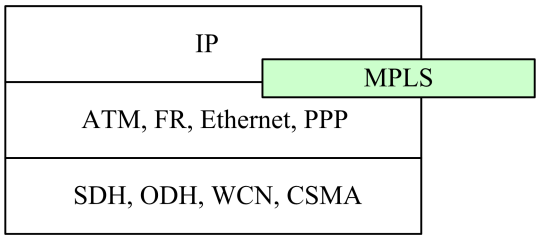
MPLS uses simplified ATM switching technology to implement switching in layer 2 and layer 3. It provides an identifier for each IP message, and encapsulates the identifier to a new MPLS message together with the IP message, deciding the transmission path and priority of the IP message.

The differences between MPLS and traditional WAN technology are in the label distribution method and label stack introduced by MPLS. In MPLS, the use of label stack brings new applications, such as traffic engineering, VPN and Fast Re-Routing (FRR).

Before forwarding the IP packet according to corresponding path, the MPLS router reads the header identifier of the message instead of the IP address in each IP message. Therefore, the speed to

forward packets is much faster, and MPLS is called a layer 2.5 protocol, as shown in Figure 3.

FIGURE 3 POSITION OF MPLS



MPLS can use different layer 2 protocols. Till now, MPLS task force has implemented standardization of identifiers used in FR, ATM, PPP link and IEEE802.3 LANs. When MPLS runs in FR and ATM, it also carries free connectivity in IP network to these connection-orientated technologies.

The working mechanism of MPLS network is that it implements routing through IP outside the MPLS network and implements switching through looking up labels within MPLS network.

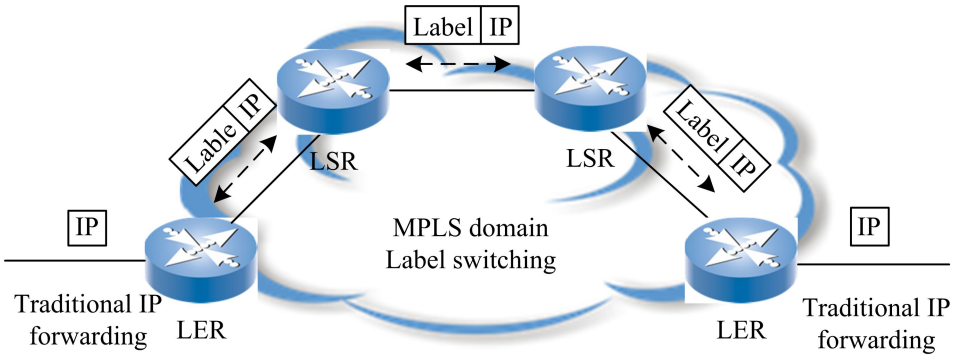
MPLS has the following advantages:

- MPLS provides connection-orientated services for IP network.
- Integrated technologies of data link layer and network layer, MPLS solves the problems about Internet extension, QoS.
- MPLS uses accurate matching instead of the longest-matching method to choose path through the short and fixed labels, which improves forwarding rate.
- When supporting IP services, MPLS also provides high reliability and QoS assurance.
- MPLS establishes CR-LSP through explicit routing function and signaling protocols with QoS parameters, therefore, traffic engineering can be implemented effectively.
- MPLS can support VPN with label nesting technology.

MPLS Working Principle

In the network outside MPLS domain, packets are forwarded in traditional IP forwarding way; while within the MPLS domain, packets are forwarded by label switching, as shown in Figure 4.

FIGURE 4 MPLS WORKING PRINCIPLE



Within the MPLS domain, all routers run the label distribution protocol, such as LDP and RSVP.

The procedure of forwarding an IP packet in the MPLS domain is described as follows:

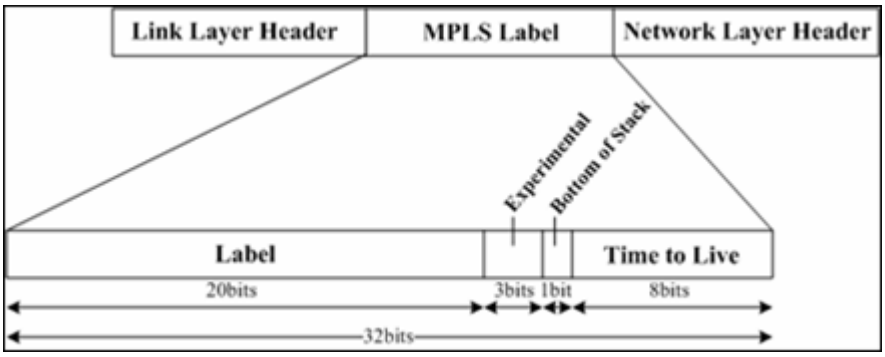
1. The ingress LSR receives the IP packet, and assigns a label to this packet for identification.
2. When the backbone LSRs receive the packet, they look up label forwarding tables, and use a new label to replace the primary label.
3. When egress LSR receives the packet, it removes the label and implements traditional IP forwarding.

MPLS Terms

MPLS Label

Label is a fixed-length and locally effective identifier. It is used to identify an FEC that a group belongs to uniquely. For example, when load balance is configured, the same FEC may correspond to multiple labels. However, only one label stands for the FEC. The structure is shown in Figure 5.

FIGURE 5 MPLS LABEL STRUCTURE



A label contains four fields:

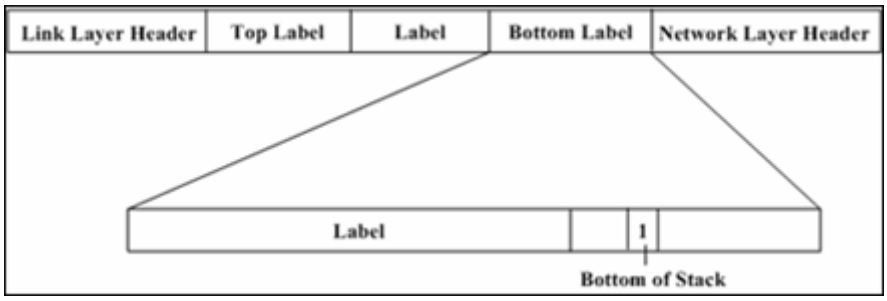
- Label field: 20 bits long, used to forward pointers.
- Exp field: 3 bits long, reserved for defining experiments.
- S field: 1 bit long. MPLS supports multiple layers of labels. When the value of this field is 1, it means that this entry is a label that locates the bottom of the label stack. When the value of this field is 0, it means that this entry is a label that locates any position of the label stack except the bottom.
- TTL field: 8 bits long, used for coding time to live.

MPLS supports different protocols in data link layer. MPLS label is encapsulated after data link layer message and before layer 3 data. Each protocol has different protocol number defined for MPLS.

In Ethernet networks, MPLS packets are identified by value 0x8847 (for unicast) and value 0x8848 (for multicast). In PPP, a new NCP, MPLSCP, is identified by value 0x8281.

The S field of a label is shown in Figure 6. In a MPLS domain, a message can be nested with multiple labels.

FIGURE 6 MPLS LABEL



When a message is encapsulated with multiple labels, LSR deals the message with First In First Out (FIFO) rule. That is, LSR just forwards the message according to the top label.

Forward Equivalence Class

Forward Equivalence Class (FEC) is a representation of a group of packets that share the same requirements for their transport. All packets in such a group are provided with the same treatment. During labeling binding of LDP, different FECs correspond to different labels. Each node in MPLS domain identifiers the FEC to which the packet belongs according to the label.

When two groups of packets with the same source address and different destination addresses enter a MPLS domain, MPLS judges the groups according to FEC and MPLS finds that the two groups belong to two FECs. Therefore, MPLS treats the groups in different ways, for example, in different paths and reserving different resources. The two groups are classified into different classes at the ingress of the MPLS domain and encapsulated different labels. The nodes in the MPLS domain forward the groups of packets according to labels. When the packets leave the MPLS domain, the

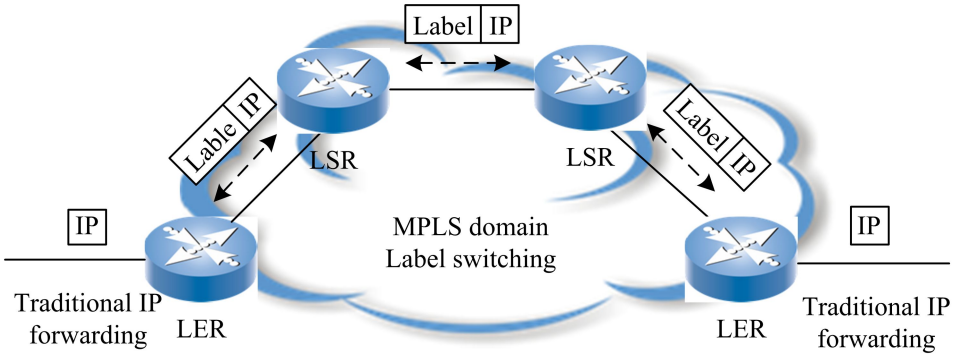
egress LSR removes the labels and forwards the packets to the destinations.

Opposed to conventional IP forwarding, in MPLS, the assignment of a particular packet to a particular FEC is done just once. It is done when the packet enters MPLS domain.

MPLS Network Structure

The structure of MPLS network is shown in Figure 7.

FIGURE 7 MPLS NETWORK STRUCTURE



Label Edge Router (LER) is at the edge of a MPLS domain. LSR assigns labels for packets entering MPLS domain. It implements functions such as traffic classification, label mapping and label removing, Label Switch Path (LSP) initiation, IP packets forwarding and Diff-Serv.

Label Switch Router (LSR) is the core router in a MPLS domain. It supports label switching and label distribution functions. It participates in establishing LSP and initiates next-hop change.

Routers in MPLS domain run Label Distribution Protocol (LDP) to distribute labels for packets.

LSP Establishment

A LSP is a forwarding path established according to MPLS. It is composed of LSRs and links from source LSR to destination LSR, which is similar to virtual circuit of ATM.

There are three drive modes to establish a LSP:

- Packet-based

  The packet traffic reaching a LSR springs label distribution. In this mode, the cost for label distribution is direct proportion to the traffic. Time delay exists for label distribution.

- Topology-based

Topology-based drive is to distribute labels for corresponding routes of routing protocols. When a LSR deals with route update of OSPF or BGP, the LSR modifies the entries in label forwarding table, meanwhile the LSR distributes labels for these entries. As long as there is a route, the label for this route is distributed. Therefore, there is no time delay when LSR forwards packets. This mode is widely used.

- Application-based

  In this mode, LSR distributes labels according to control service traffic with normal requirements. The corresponding protocol is RSVP. When a LSR deals with RSVP, the LSR modifies the entries in label forwarding table, meanwhile the LSR distributes labels for these entries. This mode requires application programs bring forward label request and stream rule in advance to obtain labels. It also distributes labels for existed routes, so there is no time delay. However, it is difficult to implement RSVP in the whole network, so this mode is seldom used.

Comparing with packet-based drive, topology-based drive has the following advantages:

- Label distribution corresponds to control message, therefore, network cost is little.
- Label is distributed before traffic arrives, so there is no time delay.

The protocols used to distribute labels for LSRs are called signalling protocols. Common signalling protocols include:

- Label Distribution Protocol (LDP)
- Constrained Route LDP (CR-LDP)
- Resource Reservation Protocol - Traffic Extension (RSVP-TE)
- Multiprotocol BGP (MP-BGP)

LSP Establishment Procedure

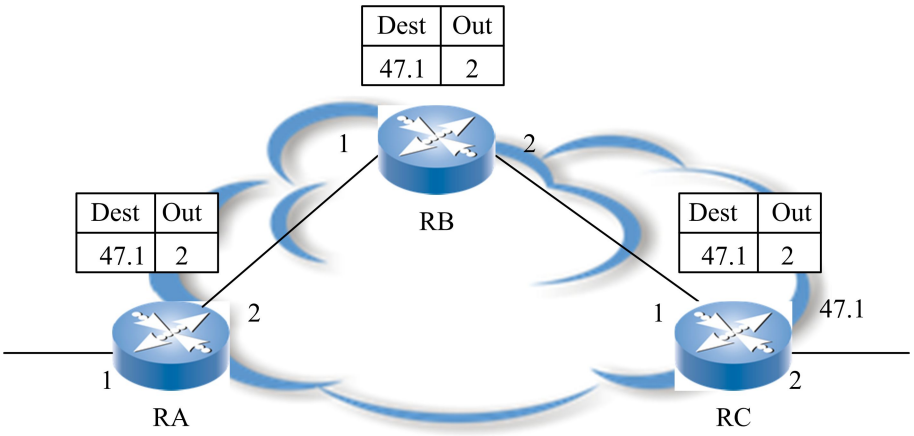
In a MPLS domain, LSP establishment procedure is described as follows:

1. Each node runs dynamic routing protocols such as BGP, OSPF and IS-IS to generate a routing table.
2. According to the routing table, each node establishes a Label Information Base (LIB) under the control of LDP.
3. **Out** labels and **in** labels from ingress LSR to egress LSR are mapped to form a LSP.

The detailed procedure is described as follows:

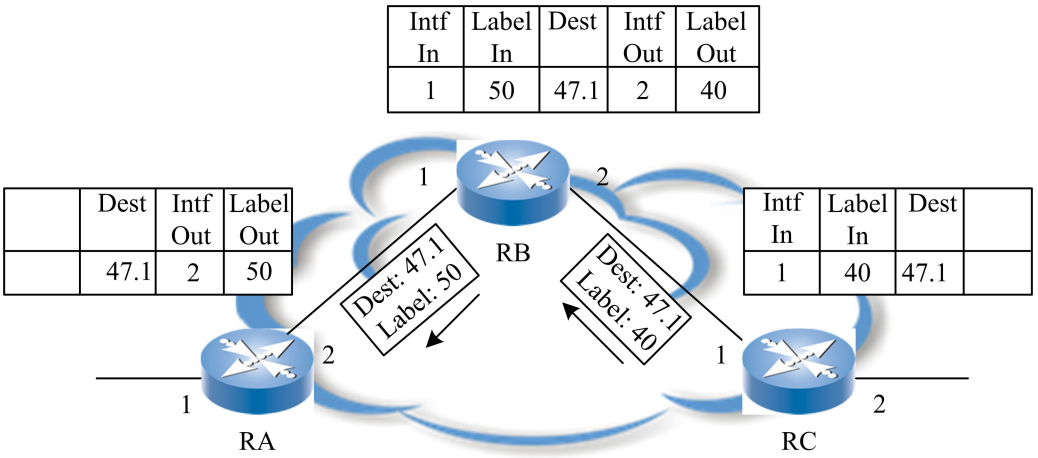
Step 1 As shown in Figure 8, each router runs OSPF to generate a routing table. RA, RB and RC learn the route 47.1.0.0/16 to LER.

FIGURE 8 GENERATING A ROUTE TABLE



Step 2 As shown in Figure 9, the routers run LDP to distribute labels.

FIGURE 9 GENERATING A LIB



As an egress LSR to 47.1.0.0/16, RC distributes a label "40" and sends it to upstream neighbor RB. This information is recorded in LIB. When RC receives a message with label "40", it knows that this message is to 47.1.0.0/16.

When RB receives the binding information of 47.1.0.0/16 and label "40" from RC, it keeps the label information and receiving interface information in LIB. Meanwhile, RB distributes a label for 47.1.0.0/16 and sends this information to neighbors except the neighbor connecting to the receiving interface. Suppose that RB sends a label "50" to RA. Therefore, there is such information in LIB of RB, as shown below:

```
IntfIn    LabelIn     Dest          IntfOut   LabelOut
1           50        47.1.0.0        2         40
```

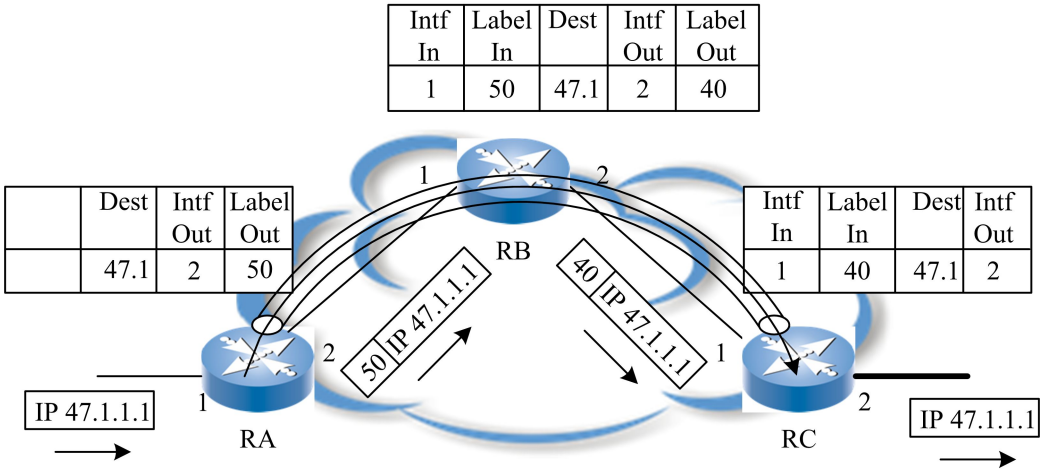
This information means, when RB receives a message with label "50" from interface 1, it changes the label to "40" and sends the message from interface 2. RB does not look up route in its routing table.

When RA receives binding information, RA acts in the same way as RB does.

LIB is always isochronous with routing tables. Once a new non-BGP route is generated in the routing table, LSR distributes a label for this route. LSR does not distribute labels for BGP routes. Instead, it distributes a label for the next hop of a BGP route.

Step 3 After the interaction of labels finishes, a LSP forms. When LSRs forward messages, they forward them according to labels instead of looking up routing tables, as shown in Figure 10.

FIGURE 10 GENERATING A LSP



When RA receives a message with destination address 47.1.1.1, it looks up its routing table first, and then it looks up label forwarding table. When it finds that the FEC 47.1.0.0/16 corresponds to the label "50", RA adds the label to the header of the message, and sends it from interface 2.

When RB receives the message with label "50" from interface 1, it looks up the label forwarding table, then RA changes the label to "40" and sends the message from interface 2.

When RC receives the message with label "40" from interface 1, it looks up the label forwarding table. RC finds that the destination of the message is a network segment connecting to itself directly. Therefore, RC removes the label and sends the IP message.

Penultimate Hop Popping

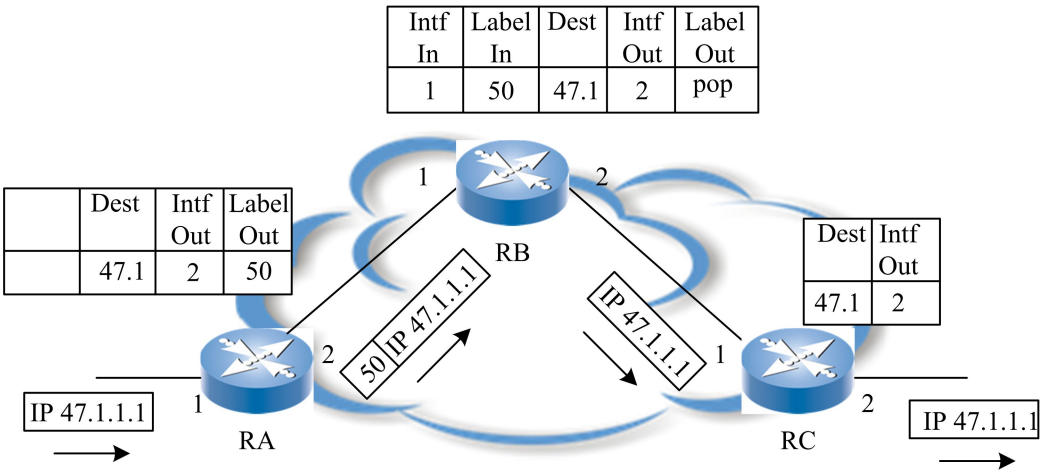
In a MPLS domain, when the egress LER receives message from its neighbor, it looks up label forwarding table first and removes the label, then it looks up routing table and forwards the IP message. This decreases the performance of the LER and increases forwarding complexity. Penultimate hop popping mechanism solves this problem.

By default, LDP enables only penultimate hop popping mechanism for direct routes and aggregated routes. For direct routes, LER has to look up routing table to get next hop information before it

forwards messages to destinations connecting to it directly. For aggregated routes, LER has to look up routing table for accurate routes. In other situations, Layer 2 information of messages is recorded in LFIB, so there is no need to look up routing table. The messages are switched directly.

As shown in Figure 11, RC is an egress LER to 47.1.0.0/16, so RC distributes a particular label "3" (means implicit-null) for messages to 47.1.0.0/16. When RB receives the label "3" distributed by RC, RB knows that it is the penultimate hop popping LSR.

FIGURE 11 PENULTIMATE HOP POPPING



During forwarding procedure, when RB receives a message with a label "50" from RA, it looks up the label forwarding table. RB finds that the "out" label is "3", so it removes the label in the message and sends the message to RC. When RC receives the message without a label, it looks up routing table directly and forwards the message.

Label Control and Management

In MPLS, there are two label distribution modes:

- Downstream-on-Demand (DoD)
- Downstream Unsolicited (DU)

In MPLS, there are two label control modes:

- Independent
- Ordered

In MPLS, there are two label holding modes:

- Liberal
- Conservative

By default, devices of ZTE use DU mode, independent mode and liberal mode.

Label Distribution Mode

In a MPLS system, a downstream LSR distributes labels for specific FEC and informs upstream LSR. That is, labels are designated by down stream and distributed from down stream to up stream.

Note:

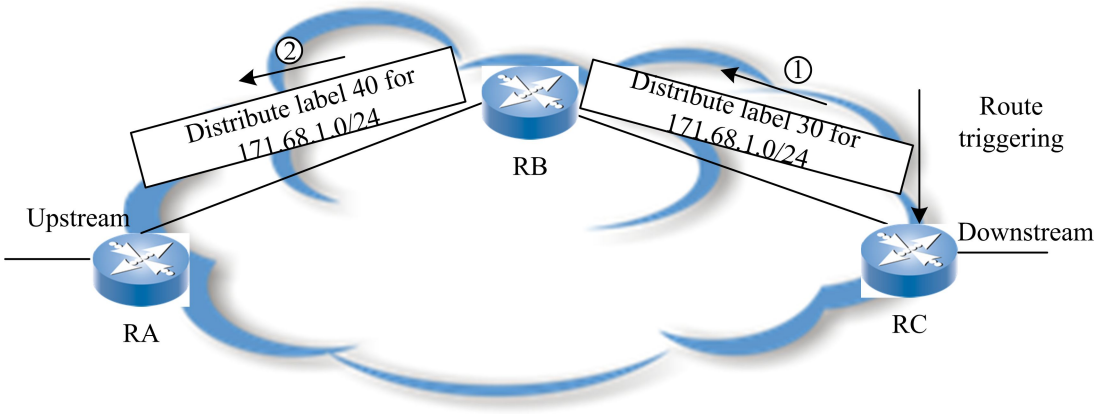
The neighbor downstream LSR and upstream LSR should come to an agreement about which label distribution mode is used. Otherwise, LSP can not be established normally.

There are two label distribution modes in MPLS: Downstream Unsolicited (DU) and Downstream on Demand (DoD).

DU For a particular FEC, a LSR does not need to get request messages for labels before designating and distributing labels.

As shown in Figure 12, RC does not need to wait for label request message from upstream to tell downstream LSR the binding information of the FEC and label. In the same way, RB does not need wait for label request message from RA to tell RC the binding information of the FEC and label.

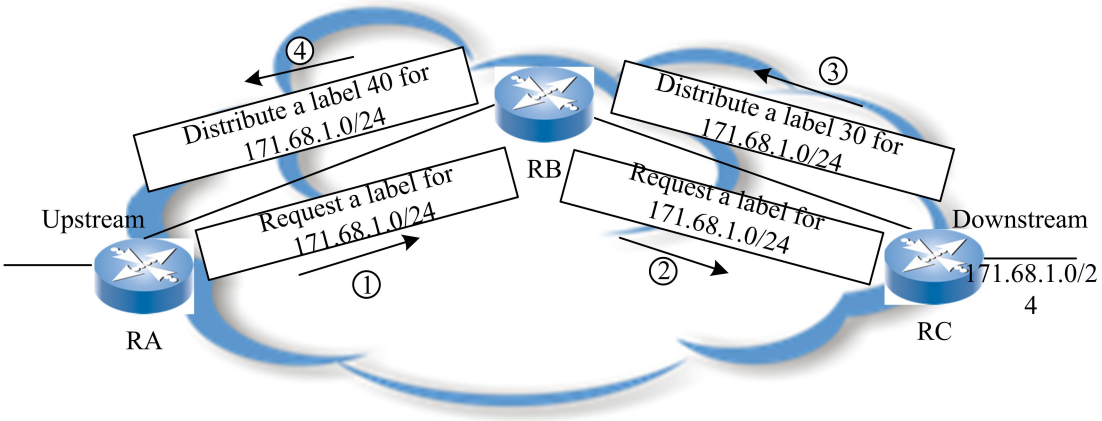
FIGURE 12 DOWNSTREAM UNSOLICITED



DoD For a particular FEC, a LSR designates and distributes labels after it gets request messages for labels.

As shown in Figure 13, RC is the egress LER to 171.68.10.0/24. RC can not send label binding information to upstream neighbor RB forwardly. RC has to wait for the request message from the upstream.

FIGURE 13 DOWNSTREAM ON DEMAND



Label Control Mode

There are two label control modes:

- Independent

  Each LSR can notify label mapping messages to neighbor LSR at any time.

- Ordered

  Only when a LSR receives specific label mapping messages of a particular FEC, or when the LSR is the egress, the LSR sends label mapping messages to upstream LSR.

Label Holding Mode

Label holding mode means the mode to process label bindings that a LSR received but currently not in use. There are two label holding modes: liberal and conservative.

Suppose there are two LSRs named LSR1 and LSR2, and LSR2 is not the next hop of LSR1. For a particular FEC, when LSR1 receives label binding from LSR2, if LSR1 saves this binding, this means that LSR1 uses liberal label holding mode. If LSR1 discards the binding, this means that LSR1 uses conservative label holding mode.

In liberal label holding mode, a LSR can adapt itself to route changes quickly. In conservative label holding mode, a LSR can reduce the number of label bindings.

Label Forwarding Table

In a label forwarding table, LSR stores binding information of FECs and labels. A LSR maintains the label forwarding through label distribution protocol dynamically. When forwarding messages, LSR looks up the table.

A label forwarding table includes the following terms:

- InLabel: "in" label, distributed by local router for upstream LSR, corresponding to FEC.
- OutLabel: "out" label, distributed by downstream LSR for local router, corresponding to FEC.
- Dest: destination network segment or host, that is, the bound FEC.
- Pfxlen: prefix length, that is, the mask of FEC.
- Interface: out interface.
- NextHop: next hop.

When a LSR receives a message, it looks up the label forwarding table. It searches the table according to InLabel of the message; then the LSR replaces the Inlabel with Outlabel and sends the message from the interface.

Label Distribution Protocol

MPLS system (RFC3031) defines the regulations of label distribution protocols. Labels are distributed automatically. Label distribution protocols are used to mark the FEC between LSRs and mark the label mapping relationships.

There are the following types of protocols implementing label distribution function:

- Label Distribution Protocol (LDP)
- Constrained Route LDP (CR-LDP)
- Resource Reservation Protocol - Traffic Extension (RSVP-TE)
- Multiprotocol BGP (MP-BGP)

LDP is a protocol that generates labels dynamically. It is based on UDP/TCP protocols. The protocol messages are routed hop by hop according to routing tables. LDP marks the FEC between LSRs, and it also marks the label mapping relationships then forms a LSP. It associates FEC with LSP, and maps FEC traffic to the LSP.

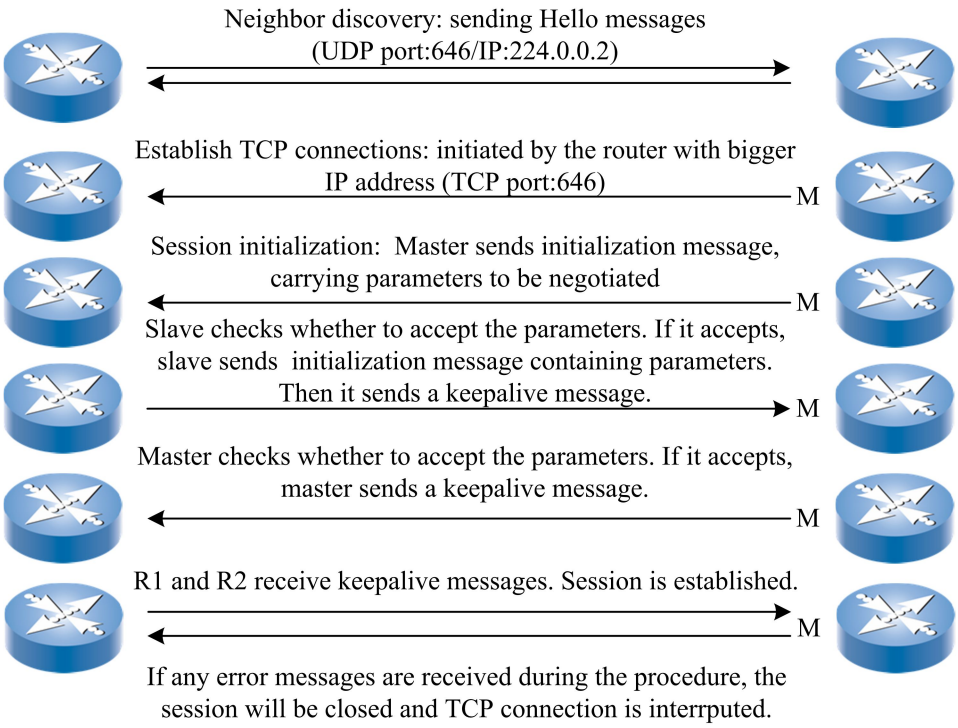
LDP Session Establishment and Maintenance

ZXR10 supports LDP regulations defined in RFC3036, including neighbor discovery, label request, label mapping, label repeal, label release and error management.

- Neighbor discovery: LSR sends Hello messages to neighbors periodically to discover LDP peers.
- Session establishment and maintenance: LSRs establish TCP connections and finish session initializations.
- LSP establishment and maintenance: LSRs distribute labels for FECs and establish LSPs.
- Session repeal: when session hold time ends, the session is interrupted.

The detailed procedure of LDP session establishment is described as follows, as shown in Figure 14.

FIGURE 14 LDP SESSION ESTABLISHMENT


Neighbor discovery: sending Hello messages
(UDP port:646/IP:224.0.0.2)
Establish TCP connections: initiated by the router with bigger IP address (TCP port:646)
M
Session initialization: Master sends initialization message, carrying parameters to be negotiated
M
Slave checks whether to accept the parameters. If it accepts, slave sends initialization message containing parameters. Then it sends a keepalive message.
M
Master checks whether to accept the parameters. If it accepts, master sends a keepalive message.
M
R1 and R2 receive keepalive messages. Session is established.
M
If any error messages are received during the procedure, the session will be closed and TCP connection is interrputed.

1. R1 and R2 send multicast Hello messages to protocol interfaces. The messages are encapsulated by UDP and port number is 646. Meanwhile, protocol interfaces on R1 and R2 monitor Hello messages in real time to discover neighbors.
2. After R1 and R2 receive Hello messages, they judge whether the session is established with the peer. If session has not been established, they prepare to establish a session. The session is a TCP connection establishment procedure, using port number 646. Before establishing the TCP connection, R1

and R2 choose a master according to IP addresses. The router with bigger IP address will acts as the master. The master initiates TCP connection. Here suppose that IP address of R2 is bigger. Therefore, R2 sends TCP connection request.

3. R2 sends an initialization message to establish session. The message contains parameters to be negotiated.
4. When R1 receives the message from R2, it checks the parameters. If parameters pass the check, R1 will send an initialization message and a keepalive message to R2. Parameters to be negotiated are included in the initialization message.
5. R2 checks the message from R1. If parameters pass the check, R2 will send an initialization message to R1.
6. The session is established. During the procedure, if any error messages are received, the session will be closed and the TCP connection is interrupted.

Configuring MPLS

Enabling MPLS

To enable MPLS, perform the following steps.

Step	Command	Function
1	`ZXR10(config)#`**mpls ip**	This enables LDP
2	`ZXR10(config)#`**interface vlan** <*vlan-number*>	This enters VLAN Layer 3 interface configuration mode
3	`ZXR10(config-if)#`**mpls ip**	This enables MPLS LDP on VLAN Layer 3 interface

Note:

To disable MPLS LDP globally, use **no mpls ip** command in global configuration mode.

To disable MPLS LDP on an interface, use **no mpls ip** command in interface configuration mode.

Configuring LDP Router-ID

To configure LDP router-ID, use the following command.

Command	Function
ZXR10(config)#**mpls ldp router-id** <*interface-name*>[**force**]	This configures LDP router-ID

In default situation, rules for selecting router ID of an LDP on ZXR10 8900 series switch are as follows:

- When **mpls ldp router-id** command is used to designate address of an interface as router ID, and an interface has an IP address and is in UP status, this interface will serve as the router ID.
- Otherwise, when there are loopback interfaces configured with an IP address, maximum IP address among the IP addresses of all loopback interfaces will serve as router ID.
- Otherwise, maximum IP address of an interface in UP status is selected as router ID.

When **force** keyword is configured, router ID designated by **mpls ldp router-id** command will come into use when next router ID selection occurs. Router ID selection occurs at the following situations:

- LDP restarts.
- Interface with address used by current router ID shuts down.
- Address of interface with address used by current router ID is not configured.

When **force** keyword is configured, effect of the command mpls ldp router-id depends on the current status of the designated interface.

- When current status of designated interface is UP, and other addresses are not the current router ID. Router will force router ID to change into designated value, intermit current session, free label studied in the session and intermit the MPLS transmission relating to binding with labels.
- When current status of the designated interface is DOWN, once it turns to UP. Router will force router ID to change into designated value, intermit current session, free label studied in the session and intermit MPLS transmission relating to binding with labels.

Configuring Label Distribution and Control

To generate and distribute labels, perform the following steps.

Step	Command	Function
1	ZXR10(config)#**mpls ldp access-fec** {**for** <*prefix-access-list*>| **host-route-only**}	This configures the labels for network segments LDP can generate
2	ZXR10(config)#**mpls label range** <*min-label*><*max-label*>	This sets usable label range of LDP
3	ZXR10(config)#**mpls ldp advertise-labels** [**for** <*prefix-access-list*>[**to** <*peer-access-list*>]]	This controls LDP distribute labels

Configuring LDP Time Parameters

To configure LDP time parameters, perform the following steps.

Step	Command	Function
1	ZXR10(config)#**mpls ldp discovery hello** {**holdtime** <*holdtime*>| **interval** <*interval*>}	This configures interval for sending LDP hello message between LSRs that connect directly and indirectly
2	ZXR10(config)#**mpls ldp holdtime** <*seconds*>	This configures the holding time when a LDP session can not receive successor LDP messages
3	ZXR10(config)#**mpls ldp backoff** <*initial-backoff*><*maximum-backoff*>	This configures the parameters of backing off re-establishing mechanism for LDP sessions

Configuring MD5 Password

To configure MD5 password, use the following command.

Command	Function
ZXR10(config)#**mpls ldp neighbor** <*ip-address*> **password** <*password*>	This configures MD5 password

Clearing LSR Peer Connection

To clear and re-establish a LSR peer connection, use the following command.

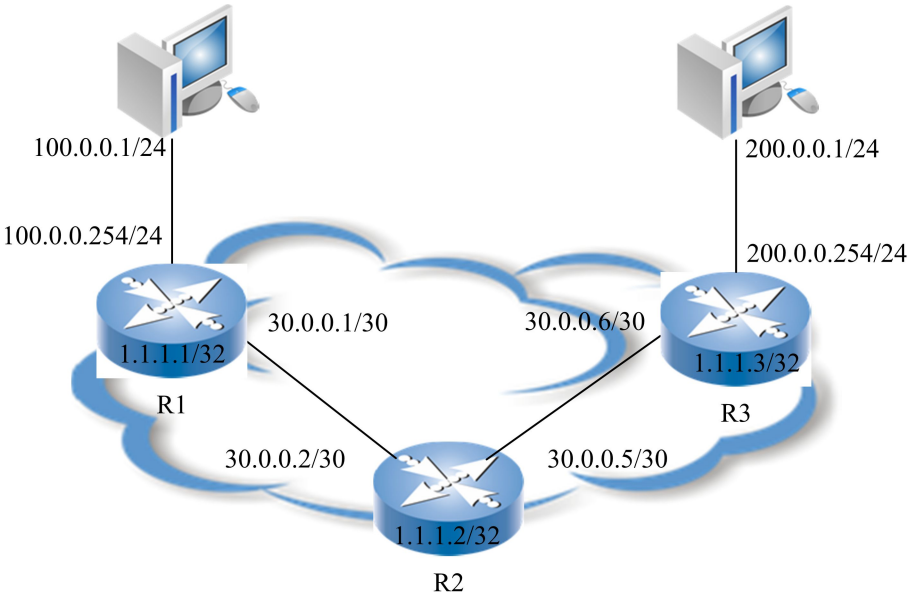
Command	Function
ZXR10(config)#**clear mpls ldp** [**neighbor** <*ip-address*>]	This clears and re-establishes a LSR peer connection

MPLS Configuration Examples

MPLS Basic Configuration Example

As shown in Figure 15, R1, R2 and R3 are in the same MPLS domain. R1 is a switch of ZXR10 8900 series switch. R2 and R3 are routers.

FIGURE 15 MPLS BASIC CONFIGURATION EXAMPLE



R1, R2 and R3 run OSPF and LDP to build a MPLS domain.

Configuration on R1:

```
R1(config)#interface loopback1
R1(config-if)#ip address 1.1.1.1 255.255.255.255
R1(config-if)#exit
R1(config)#mpls ldp router-id loopback1 force
R1(config)#vlan 10
R1(config-vlan)#switchport pvid gei_1/1
R1(config-vlan)#exit
R1(config)#interface vlan 10
R1(config-if)#ip address 30.0.0.1 255.255.255.252
R1(config-if)#mpls ip
R1(config-if)#exit
R1(config)#vlan 20
R1(config-vlan)#switchport pvid gei_1/2
R1(config-vlan)#exit
R1(config)#interface vlan 20
```

```
R1(config-if)#ip address 100.0.0.254 255.255.255.0
R1(config-if)#exit
R1(config)#router ospf 1
R1(config-router)#network 1.1.1.1 0.0.0.0 area 0
R1(config-router)#network 30.0.0.1 0.0.0.3 area 0
R1(config-router)#redistribute connected
R1(config-router)#exit
R1(config)#mpls ip
```

Configuration on R2:

```
R2(config)#interface loopback1
R2(config-if)#ip address 1.1.1.2 255.255.255.255
R2(config-if)#exit
R2(config)#mpls ldp router-id loopback1 force
R2(config)#interface fei_0/1
R2(config-if)#ip address 30.0.0.2 255.255.255.252
R2(config-if)#mpls ip
R2(config-if)#exit
R2(config)#interface fei_0/2
R2(config-if)#ip address 30.0.0.5 255.255.255.252
R2(config-if)#mpls ip
R2(config-if)#exit
R2(config)#router ospf 1
R2(config-router)#network 1.1.1.2 0.0.0.0 area 0
R2(config-router)#network 30.0.0.2 0.0.0.3 area 0
R2(config-router)#network 30.0.0.5 0.0.0.3 area 0
R2(config-router)#exit
R2(config)#mpls ip
```

Configuration on R3:

```
R3(config)#interface loopback1
R3(config-if)#ip address 1.1.1.3 255.255.255.255
R3(config-if)#exit
R3(config)#mpls ldp router-id loopback1 force
R3(config)#interface fei_0/1
R3(config-if)#ip address 30.0.0.6 255.255.255.252
R3(config-if)#mpls ip
R3(config-if)#exit
R3(config)#interface fei_2/1
R3(config-if)#ip address 200.0.0.254 255.255.255.0
R3(config-if)#exit
R3(config)#router ospf 1
R3(config-router)#network 1.1.1.3 0.0.0.0 area 0
R3(config-router)#network 30.0.0.6 0.0.0.3 area 0
R3(config-router)#redistribute connected
R3(config-router)#exit
R3(config)#mpls ip
```

The information on interface of R1 which enables MPLS is shown below.

```
R1#show mpls interface
interface of LDP:
Interface            IP        Tunnel   Operational
vlan10            Yes(ldp)     No       Yes
```

Yes means the startup is normal.

LDP session information on R1 is shown below.

```
R1#show mpls ldp neighbor
Peer LDP Ident: 1.1.1.2:0; Local LDP Ident 1.1.1.1:0
        TCP connection: 1.1.1.2.1024 - 1.1.1.1.646
        state: Oper; Msgs sent/rcvd: 26/25; Downstream
        Up Time: 00:16:07
        LDP discovery sources:
          vlan10; Src IP addr: 30.0.0.2
        Addresses bound to peer LDP Ident:
          1.1.1.2  30.0.0.2  30.0.0.5
```

The above information shows that a TCP connection is established normally between LSRs. If a TCP connection is not established normally, the information is shown as below.

```
R1#show mpls ldp neighbor
Peer LDP Ident: 1.1.1.2:0; Local LDP Ident 1.1.1.1:0
        No TCP connection
        state: Non; Msgs sent/rcvd: 0/0; Downstream
        Up Time: 00:01:46
        LDP discovery sources:
          vlan10; Src IP addr: 30.0.0.2
        Addresses bound to peer LDP Ident:
```

LIB information on R1 is shown below.

```
R1#show mpls ldp bindings
 1.1.1.1/32
        local binding:  label: imp-null
        remote binding: lsr: 1.1.1.2:0, label: 18
 1.1.1.2/32
        local binding:  label: 18
        remote binding: lsr: 1.1.1.2:0, label: imp-null(inuse)
 1.1.1.3/32
        local binding:  label: 17
        remote binding: lsr: 1.1.1.2:0, label: 16(inuse)
 30.0.0.0/30
        local binding:  label: imp-null
        remote binding: lsr: 1.1.1.2:0, label: imp-null
 30.0.0.4/30
        local binding:  label: 16
        remote binding: lsr: 1.1.1.2:0, label: imp-null(inuse)
 100.0.0.0/24
        local binding:  label: imp-null
        remote binding: lsr: 1.1.1.2:0, label: 19
 200.0.0.0/24
        local binding:  label: 19
        remote binding: lsr: 1.1.1.2:0, label: 17(inuse)
```

Local binding means local label distribution and notifying to other LSRs. **Remote binding** means that the label is notified by LSP peer. For local network segment, when a LSR receives the label **imp-null**, the LSR implements penultimate hop popping mechanism.

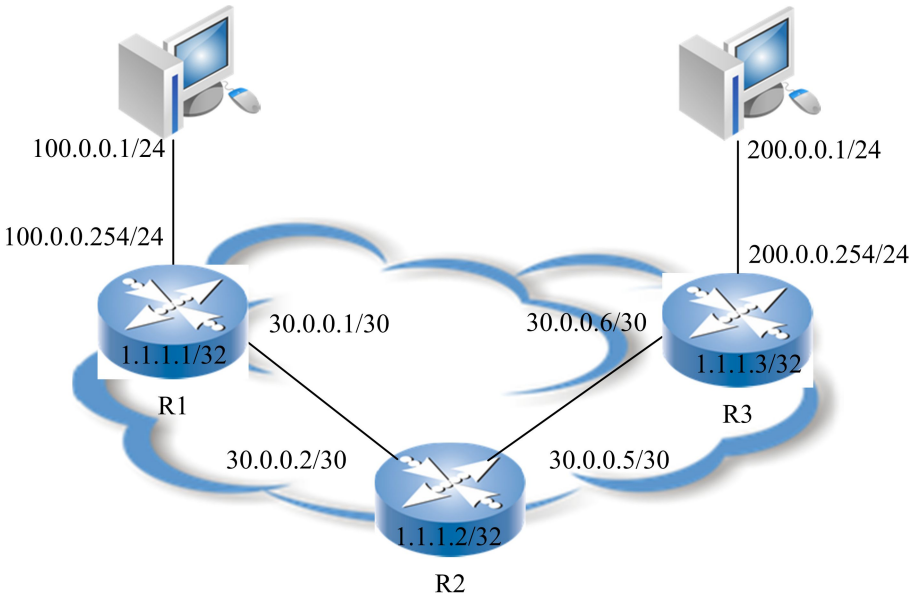
Label forwarding table information on R1 is shown below.

```
R1#show mpls forwarding-table
Mpls Ldp Forwarding-table:
InLabel   OutLabel    Dest       Pfxlen  Interface     NextHop
18        Pop tag    1.1.1.2     32      vlan10        30.0.0.2
17        16         1.1.1.3     32      vlan10        30.0.0.2
16        Pop tag    30.0.0.4    30      vlan10        30.0.0.2
19        17         200.0.0.0   24      vlan10        30.0.0.2
```

Label Distribution Example for Specified Nerwork Segment

As shown in Figure 16, R1, R2 and R3 are in the same MPLS domain. R1 is a switch of G series. R2 and R3 are routers.

FIGURE 16 LABEL DISTRIBUTION EXAMPLE



In MPLS VPN, LDP only generates labels for MPLS Router-ID.

LIB information on R1 is shown below.

```
R1#show mpls ldp bindings
 1.1.1.1/32
        local binding:  label: imp-null
        remote binding: lsr: 1.1.1.2:0, label: 19
 1.1.1.2/32
        local binding:  label: 18
        remote binding: lsr: 1.1.1.2:0, label: imp-null(inuse)
 1.1.1.3/32
        local binding:  label: 17
        remote binding: lsr: 1.1.1.2:0, label: 16(inuse)
 30.0.0.0/30
        local binding:  label: imp-null
        remote binding: lsr: 1.1.1.2:0, label: imp-null
 30.0.0.4/30
        local binding:  label: 16
        remote binding: lsr: 1.1.1.2:0, label: imp-null(inuse)
 100.0.0.0/24
        local binding:  label: imp-null
        remote binding: lsr: 1.1.1.2:0, label: 18
 200.0.0.0/24
        local binding:  label: 19
        remote binding: lsr: 1.1.1.2:0, label: 17(inuse)
```

Configuration on R1:

```
R1#configure terminal
R1(config)#acl standard number 1
R1(config-std-acl)#rule 1 permit 1.1.1.1 0.0.0.0
R1(config-std-acl)#rule 2 permit 1.1.1.2 0.0.0.0
R1(config-std-acl)#rule 3 permit 1.1.1.3 0.0.0.0
R1(config-std-acl)#exit
R1(config)#mpls ldp access-fec for 1
R1(config)#mpls ldp access-fec force
```

Now LIB information on R1 is shown as follows.

```
R1#sh mpls ldp bindings
 1.1.1.1/32
        local binding:  label: imp-null
        remote binding: lsr: 1.1.1.2:0, label: 19
 1.1.1.2/32
```

```
        local binding:  label: 18
        remote binding: lsr: 1.1.1.2:0, label: imp-null(inuse)
 1.1.1.3/32
        local binding:  label: 17
        remote binding: lsr: 1.1.1.2:0, label: 16(inuse)
```

The above information shows that R1 only generates labels for router-ID.

Configurations on R2 and R3 are similar with that on R1.

MPLS Maintenance and Diagnosis

To view MPLS configuration information, use the following commands.

To view enabled MPLS interfaces, use the following command.

Command	Function
ZXR10#**show mpls interface**	This displays the enabled MPLS interfaces

This example describes how to view enabled MPLS interfaces of R2. **Yes** means the startup is normal.

```
ZXR10_R2#show mpls interface
interface of LDP:
Interface            IP        Tunnel Operational
VLAN10             Yes(ldp)    No     Yes
VLAN20             Yes(ldp)    No     Yes
```

To view MPLS LDP parameters, mainly LDP timer parameter, use the following command.

Command	Function
ZXR10#**show mpls ldp parameters**	This displays MPLS LDP parameters, mainly LDP timer parameter

This example displays LDP parameters message of R2.

```
ZXR10_R2#show mpls ldp parameters
Protocol version: 1
Downstream label pool: min label: 16; max label: 1048575
Session hold time: 180 sec; keep alive interval: 60 sec
Discovery hello: holdtime: 15 sec; interval: 5 sec
Downstream on Demand max hop count: 255
LDP initial/maximum backoff: 15/120 sec
LDP loop detection: off
```

To view LDP discovery message, use the following command.

Command	Function
ZXR10#**show mpls ldp discovery**	This displays the LDP discovery message

This command is to show IP address (Transport IP address) used to set up TCP link by LDP neighbors discovered on the interfaces. To set up the session, router must get the reachable route to the router, that is, it can ping address. "xmit/recv" means sending/receiving hello message on the interface, and they are indispensable.

This example shows the detailed LDP discovery information of R2.

```
ZXR10_R2#show mpls ldp discovery detail
Local LDP Identifier:
    10.10.2.2:0
Discovery Sources:
    Interfaces:
    VLAN10 (ldp): xmit/recv
    LDP Id: 10.10.1.1:0
    Src IP addr: 10.10.12.1; Transport IP addr: 10.10.12.1
    VLAN20(ldp): xmit/recv
    LDP Id: 10.10.3.3:0
    Src IP addr: 10.10.23.3; Transport IP addr: 10.10.3.3
```

To view LDP session, use the following command.

Command	Function
ZXR10#**show mpls ldp neighbor**	This displays the LDP session

This example shows the LDP session information of R2. It must set up LDP session between LSRs before distributing the labels.

```
ZXR10_R2#show mpls ldp neighbor detail
Peer LDP Ident: 10.10.1.1:0; Local LDP Ident 10.10.2.2:0
        TCP connection: 10.10.12.1.1025 - 10.10.2.2.646
        state: Oper; Msgs sent/rcvd: 240/240; Downstream
        Up Time: 03:52:25
        LDP discovery sources:
          Vlan10; Src IP addr: 10.10.12.1
            holdtime: 15000 ms, hello interval: 5000 ms
        Addresses bound to peer LDP Ident:
          10.10.12.1  10.10.1.1
        Peer holdtime: 180000 ms; KA interval: 60000 ms
```

The message above indicates that there is a normal TCP link between LDPs, with source IP address, destination IP address and port numbers of the link. When there is no normal TCP link session, it will show as follows.

```
ZXR10_R2#show mpls ldp neighbor
Peer LDP Ident: 10.10.1.1:0; Local LDP Ident 10.10.2.2:0
        No TCP connection
        state: Non; Msgs sent/rcvd: 0/0; Downstream
        Up Time: 00:00:45
        LDP discovery sources:
          Vlan10; Src IP addr: 10.10.12.1
        Addresses bound to peer LDP Ident:
```

After setting up a normal LDP session use the following command to view the LDP label bindings that have been studied.

Command	Function
ZXR10#**show mpls ldp bindings**	This displays LDP bindings that have been studied

This example shows LDP label bindings that have been studied by R2.

```
ZXR10_R2#show mpls ldp bindings
 10.10.1.1/255.255.255.255
        local binding:  label: 17
        remote binding: lsr: 10.10.3.3:0, label: 18
        remote binding: lsr: 10.10.1.1:0, label: imp-null(inuse)
 10.10.2.2/255.255.255.255
        local binding:  label: imp-null
        remote binding: lsr: 10.10.3.3:0, label: 17
        remote binding: lsr: 10.10.1.1:0, label: 18
 10.10.3.3/255.255.255.255
        local binding:  label: 16
        remote binding: lsr: 10.10.3.3:0, label: imp-null(inuse)
        remote binding: lsr: 10.10.1.1:0, label: 17
 10.10.12.0/255.255.255.0
        local binding:  label: imp-null
        remote binding: lsr: 10.10.3.3:0, label: 16
        remote binding: lsr: 10.10.1.1:0, label: imp-null
 10.10.23.0/255.255.255.0
        local binding:  label: imp-null
        remote binding: lsr: 10.10.3.3:0, label: imp-null
        remote binding: lsr: 10.10.1.1:0, label: 16:
```

Local binding means local label distribution, and notifies towards up stream. **Remote binding** means label message notified from down stream LSR, of which, for local network segment, the label is distributed as **imp-null**. The receiver discards the label at the last but one router.

Similar commands are shown in the following example.

```
ZXR10_R1#show mpls forwarding-table
Mpls Ldp Forwarding-table:
InLabel   OutLabel   Dest       Pfxlen  Interface   NextHop
18        Pop tag   10.10.2.2     32      vlan10   0.10.12.2
17         16       10.10.3.3     32      vlan10   10.10.12.2
16        Pop tag   10.10.23.0    24      vlan10   10.10.12.2

ZXR10_R2#show mpls forwarding-table
Mpls Ldp Forwarding-table:
InLabel   OutLabel   Dest        Pfxlen  Interface   NextHop
17        Pop tag   10.10.1.1      32      vlan10   10.10.12.1
16        Pop tag   10.10.3.3      32      vlan20   10.10.23.3

ZXR10_R3#show mpls forwarding-table
Mpls Ldp Forwarding-table:
InLabel  OutLabel   Dest         Pfxlen  Interface   NextHop
18         17      10.10.1.1       32      vlan20   10.10.23.2
17        Pop tag  10.10.2.2       32      vlan20   10.10.23.2
16        Pop tag  10.10.12.0      24      vlan20   10.10.23.2
```

- **InLabel** means label bundled locally.
- **OutLabel** means label studied from down stream.
- If the label notified from down stream is **imp-null**, it will do **Pop tag**.

To view MPLS debugging information, perform the following steps.

Step	Command	Function
1	ZXR10#**debug mpls ldp transport**	This monitors the message LDP discovered
2	ZXR10#**debug mpls ldp session**	This monitors the LDP session
3	ZXR10#**debug mpls ldp messages**	This monitors the message sending to/receiving from LDP neighbor
4	ZXR10#**debug mpls ldp bindings**	This monitors the address and label notified from LDP neighbor
5	ZXR10#**debug mpls ldp advertisements**	This monitors the address and label notifying to LDP neighbor

This example describes events related to mechanism discovered by LDP on R1 are monitored.

```
ZXR10_R1#debug mpls ldp transport events
LDP transport events debugging is on
ZXR10_R1#
ldp: Send ldp hello;
Vlan10, scr/dst 10.10.12.1(0.0.0.0)/224.0.0.2, intf_id 257
ldp: Rcvd ldp hello;
Vlan10, from 10.10.12.2(10.10.2.2:0), intf_id 257
ZXR10_R1#debug mpls ldp transport connections
LDP transport connection debugging is on
ZXR10_R1#
ldp: Hold timer expired for adj 0, will close adj
ldp: Closing ldp conn; 10.10.12.1:1025<-->10.10.2.2:646
ldp: Opening ldp conn; 10.10.12.1<-->10.10.2.2
ldp: Opening ldp conn; 10.10.12.1<-->10.10.2.2
ldp: ldp conn closed; 10.10.12.1:1026<-->10.10.2.2:646
ldp: ldp conn closed; 10.10.12.1:1027<-->10.10.2.2:646
ldp: Opening ldp conn; 10.10.12.1<-->10.10.2.2
ldp: ldp conn is up; 10.10.12.1:1028<-->10.10.2.2:646
```

Chapter 3

MPLS L3VPN Configuration

Table of Contents

MPLS L3VPN Overview 29
Configuring MPLS L3VPN 35
Configuring MPLS L3VPN Load Balance 45
MPLS L3VPN Configuration Example 45
Configuration Example of MPLS L3VPN Load Balance on Smartgroup Link 49

MPLS L3VPN Overview

MPLS VPN is an MPLS-based IP VPN. It is a routing method of applying the MPLS technology to network routing and switching equipment to simplify core routers. MPLS VPN uses the label switching combined with traditional routing technologies to implement IP-based VPN.

MPLS VPN can be used to construct broadband Intranet and Extranet and can meet multiple flexible service requirements.

MPLS VPN can utilize the powerful transmission capability of a common backbone network, reduce the construction costs of the Intranet, greatly improve the operation and management flexibility of user's networks, and meanwhile can meet the requirements of users for secure, realtime, broadband and convenient information transmission.

MPLS VPN Advantages

In an IP-based network, MPLS has the following advantages:

- Reducing costs

  MPLS simplifies the integration technology of ATM and IP, efficiently combines the Layer 2 and Layer 3 technologies, reduces costs and protects user's investment at earlier stages.

- Improving resource utilization

  Since label switching is used on the network, user's LANs at different points can use repeated IP addresses to improve the utilization of IP resources.

- Improving network speed

Since label switching is used, the address search time in each hop process is shortened, the transmission time of data on a network is reduced, and the network speed is improved.

- Improving flexibility and scalability

  Since MPLS uses AnyToAny connection, the network flexibility and scalability are improved. With respect to flexibility, special control policy can be customized to meet special requirements of different users and implement value-added services.

  The scalability includes the following two aspects: more VPNs on a network and easy user expansion in the same VPN.

- Convenience in user's application

  The MPLS technology will find wider application in networks of different carriers, so that an enterprise user can set up a global VPN conveniently.

- Improving security

  MPLS serves as a channel mechanism to implement transparent message transmission. LSPs of MPLS have high reliability and security similar to frame relay and ATM Virtual Channel Connection (VCC).

- Enhancing service integration capability

  A network can support the integration of data, audio and video services.

- QoS guarantee of MPLS

  Related standards and drafts drawn by IETF for BGP/MPLS VPN:

  - RFC 2547, BGP/MPLS VPN
  - Draft RFC 2547bis, BGP/MPLS VPN
  - RFC 2283, multi-protocol extension BGP4

Related Terms

A BGP/MPLS VPN network system covers the following types of network devices:

- PE

  A Provider Edge (PE) refers to a router connected to a CE in a client site on a carrier's network. A PE router supports VPN and labeling function (the labeling function can be provided by RSVP, LDP or CR-LDP).

- P

  Provider (P) router is a router in the core of network, which is not connected to any router in any customer site, but is a part of the tunnel in a PE pair. P router supports MPLS LSP or LDP, but does not need to support VPN.

- CE

  Customer Edge (CE) router, which connects to the PE router on the network, is on the network of ISP. A CE router must connect with a PE router.

VPN-IPv4 Address and Route Distinguisher

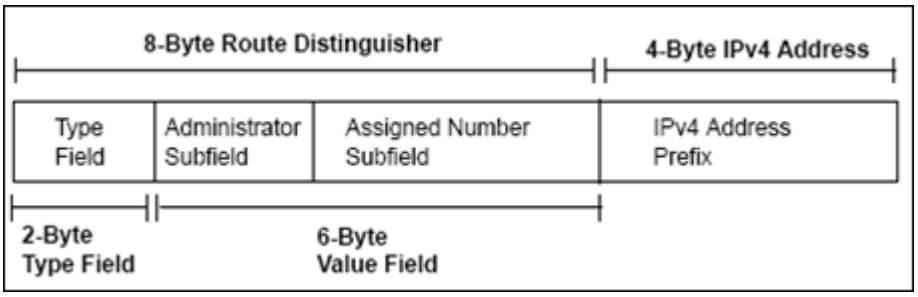
Layer 3 VPN may be connected to private networks through Internet, these private networks can use public or private addresses. When private networks use private addresses, the addresses between different private networks may be repeated.

To avoid repetition of private addresses, public addresses can be used in network equipment to replace private addresses. Solutions are provided in RFC2547bis and it uses present private net ID to create a new definite address.

This new address is one of the components in VPN-IPv4 address family and is BGP address family of MP-BGP. In a VPN-IPv4 address, there is a value used to differentiate different VPNs, called Route Distinguisher (RD).

A VPN-IPv4 address is a 12-byte quantity composed of an 8-byte Route Distinguisher followed by a 4-byte IPv4 address prefix. The structure of VPN-IPv4 address is shown in Figure 17.

FIGURE 17 STRUCTURE OF VPN-IPV4 ADDRESS



8-byte RD is composed of a 2-byte Type field and a 6-byte Value field. Type field determines the lengths of the Value field's two subfields (Administrator and Assigned Number), as well as semantics of Administrator field. The functions of the fields are described as follows:

- Type filed

  It determines the lengths of other two domains.

  If the value of the type field is 0, the administrator (ADM) subfield is four bytes and the Assignment Number (AN) subfield is two bytes.

  If the value of the type field is 1, the administrator (ADM) subfield is two bytes and the Assignment Number (AN) subfield is four bytes.

- Administrator subfield

  It identifies an administration allocation number.

  If the value of the type field is 0, the administrator subfield contains an IPv4 address. RFC2547bis recommends that the IP address of a router (this address is normally configured as router ID) should be used, and this address is a public address.

If the value of the type field is 1, the administrator subfield contains an AS number. RFC2547bis recommends that the public AS number for IANA allocation should be the AS number of the ISP or the customer.

- Assignment number subfield

  It is a number assigned by the ISPs.

  If the type field is 0, the length of the Assignment number subfield is two bytes.

  If the type field is 1, the length of the Assignment number subfield is four bytes.

An RD is only used between PEs to differentiate IPv4 addresses of different VPNs. The ingress generates an RD and converts the received IPv4 route of the CE into a VPN-IPv4 address. Before advertising the route to the CE, the egress PE converts the VPN-IPv4 route into an IPv4 route.

MPLS L3VPN Working Principles

Basic operation mode of MPLS VPN is the application of L3 technologies. Each VPN has an independent VPN-ID, users of each VPN can only communicate with members in the same VPN and only VPN members can enter VPN.

On MPLS-based VPNs, service provider assigns a distinguisher to each VPN, called Route Distinguisher (RD). Distinguisher is unique in network of service provider.

Forwarding table contains a unique address, called VPN-IP address, which is formed through connection of the RD and IP address of the user. VPN-IP address is a unique one in network. Address table is stored in forwarding table.

BGP is a routing information distribution protocol, which uses multi-protocol extension and common attributes to define VPN connectivity. On MPLS-based VPNs, BGP only advertise information to members in the same VPN and provide basic security by means of traffic split.

Data is forwarded by using LSP. LSP defines a special path that cannot be changed, to guarantee the security. Such a label-based mode can provide confidentiality as frame relay and ATM. Service provider relates a special VPN to an interface, and packet forwarding depends upon ingress labels.

VPN forwarding table contains a label corresponding to the VPN-IP address. Label is used to send data to the corresponding destination. As label is used instead of IP address, a user can maintain its dedicated address structure, without the need of data transfer by means of Network Address Translation (NAT). According to data ingress, corresponding router will select a special VPN forwarding table that only contains a valid destination address in VPN.

First, CE provides routing information of client in net to PE router by static routing, default routing or protocols like RIP, OSPF, IS-IS or BGP.

Meanwhile, CE transmits information of VPN-IP and corresponding labels (labels on VPN, shortened into internal-layer label as follows) by adopting BGP between PEs.

Traditional IGP is adopted to learn routing information from each other between PE and P routers. LDP is adopted to bind routing information and labels (labels in the backbone network, shortened as external-layer label as follows).

In this case, basic network topology and routing information of CE, PE and P routers have already been formed. A PE router has routing information of the backbone network and routing information of each VPN.

When a CE user on a VPN enters the network, system can identify to which VPN CE belongs on the interface between CE and PE, and will further read the next-hop address information in the routing table of the VPN. In addition, forwarded packets will be marked with a VPN label (internal layer label). In this case, the next-hop address obtained is the address of a PE that is the peer of this PE.

To reach the destination PE, the routing information of backbone network should be read from the source PE to obtain the address of the next P router, and meanwhile, forwarded user's packets will be tagged with a backbone network label (external layer label).

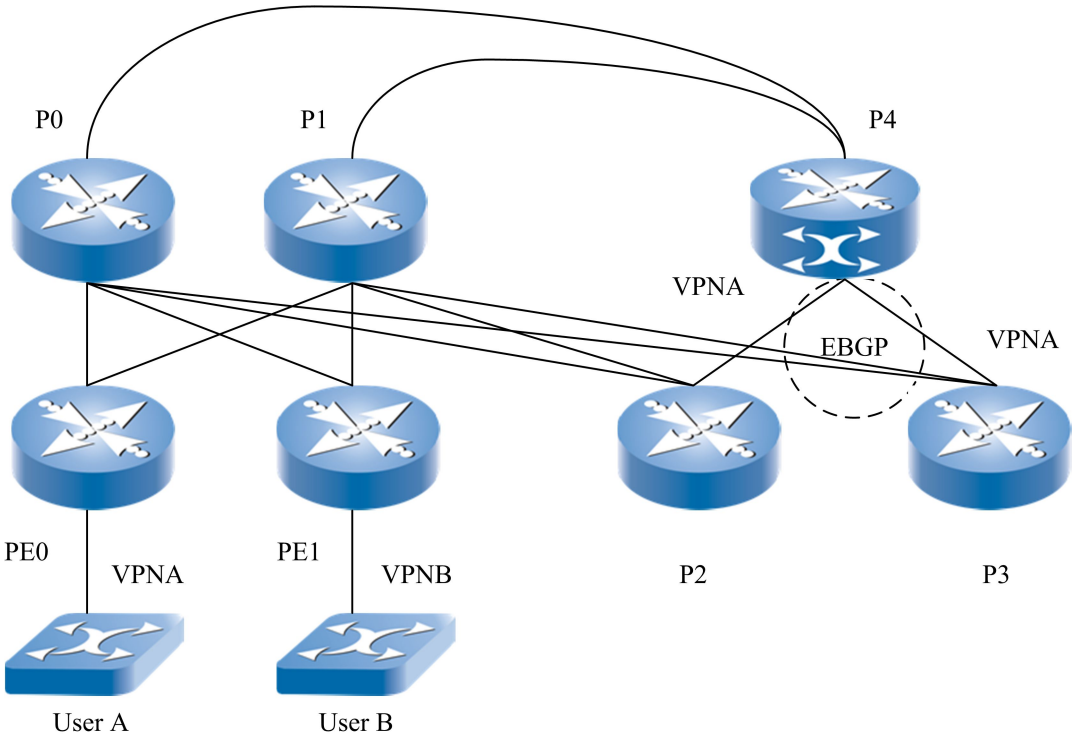
In backbone network, all P routers after the source PE read the external layer label to determine the next hop. Therefore, only simple label switching is performed on the backbone network.

When a packet reaches the last P router before arriving at the destination PE, external layer label will be cancelled. After packet reaches the destination PE, PE will read internal layer label, find the next-hop CE in the corresponding VRF, send packet to related interface and further transfer data to CE network of VPN.

MPLS L3VPN Load Balance

MPLS L3VPN load balance is to implement load sharing and protect links when switches access to VPN through PE. As shown in Figure 18, all VPN service traffic (for example, VPN traffic from P0 and P1) goes to P4 through P2 and P3. After filtration, service traffic goes back to P0 and P1 from P4 and then goes to Internet. For VPN A, traffic from user A and user B to P2 and P3, and traffic going back to P2 and P3 from P4 are load balanced.

FIGURE 18 MPLS L3VPN LOAD BALANCE



ZXR10 8900 series switch support load balance traffic forwarded by VPN, including the following types.

- Load balance traffic forwarded according to inner label
- Load balance traffic forwarded according to outer label
- Load balance traffic forwarded according to common route

MPLS L3VPN Load Balance on Smartgroup Link

In current network, services require more and more bandwidth. A common situation is that a gigabit or 10–gigabit link can not meet the requirement of services. In this situation, ECMP and aggregation of multiple links can solve the problem in Layer 3. Aggregation of multiple physical links can solve the problem in Layer 2.

MPLS/VPN function is used widely in internal metropolitan core networks and enterprise networks. In MPLS/VPN network, load balance on Smartgroup link has important function. Smartgroup link may exist between CE and PE, PE and PE, PE and P, P and P. ZXR10 8900 series switch supports MPLS L3VPN load balance on Smartgroup link. Moreover, ZXR10 8900 series switch supports different load balance policies in different situations.

ZXR10 8900 series switch can implement load balance on Smart-group link composed of GE, 10–GE and FE in the following modes.

- According to IP address (source IP, destination IP and source+destination IP)
- According to label (public network label, private network, public+private network label)

ZXR10 8900 series switch can support up to 32 Smartgroups. There are up to 8 ports in each Smartgroup. Balance efficiency ratio can reach 100%. The convergence time is within 600ms.

Configuring MPLS L3VPN

Configuring OSPF in MPLS Domain

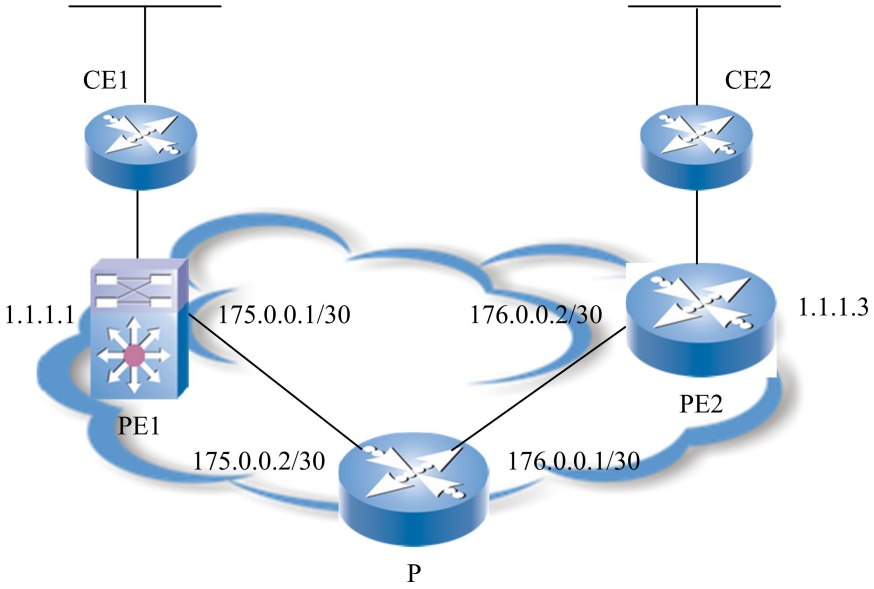
To configure OSPF in a MPLS domain, perform the following steps.

Step	Command	Function
1	`ZXR10(config)#`**router ospf** *<process-id>*	This starts OSPF process
2	`ZXR10(config-router)#`**network** *<network-number><wild-card>* **area** *<area-id>*	This enables OSPF on an interface

Example This example describes how to configure OSPF as IGP in a MPLS domain.

As shown in Figure 19, configure OSPF on PE1, P and PE2.

FIGURE 19 CONFIGURING OSPF IN MPLS DOMAIN



Configuration on PE1:

```
PE1(config)#router ospf 1
PE1(config-router)#network 1.1.1.1 0.0.0.0 area 0.0.0.0
PE1(config-router)#network 175.0.0.0 0.0.0.3 area 0.0.0.0
```

Configuration on P:

```
P(config)#router ospf 1
P(config-router)#network 175.0.0.0 0.0.0.3 area 0.0.0.0
P(config-router)#network 176.0.0.0 0.0.0.3 area 0.0.0.0
```

Configuration on PE2:

```
PE2(config)#router ospf 1
PE2(config-router)#network 1.1.1.3 0.0.0.0 area 0.0.0.0
PE2(config-router)#network 176.0.0.0 0.0.0.3 area 0.0.0.0
```

Configuring VRF on PE

A VRF is used to relate information of a VPN. Different VRFs on a PE are independent. A VRF includes an interface table, a routing table, a label forwarding table and routing protocol information.

To create a VRF and configure related parameters, perform the following steps.

Step	Command	Function
1	ZXR10(config)#**ip vrf** <*vrf-name*>	This creates a VRF
2	ZXR10(config-vrf)#**rd** <*route-distinguisher*>	This configures RD
3	ZXR10(config-vrf)#**route-target** [**import**|**export**|**both**]<*extended-community*>	This configures router target

Note:

RD is unique on local router and in the whole MPLS domain. Multiple router targets can be configured on a router.

To associate VRF with an interface, perform the following steps.

Step	Command	Function
4	ZXR10(config)#**interface vlan** <*vlan-number*>	This enters Layer 3 interface configuration mode
5	ZXR10(config-if)#**ip vrf forwarding** <*vrf-name*>	This associates VRF with an interface
6	ZXR10(config-if)#**ip address** <*ip-address*><*netmask*>	This configures an IP address

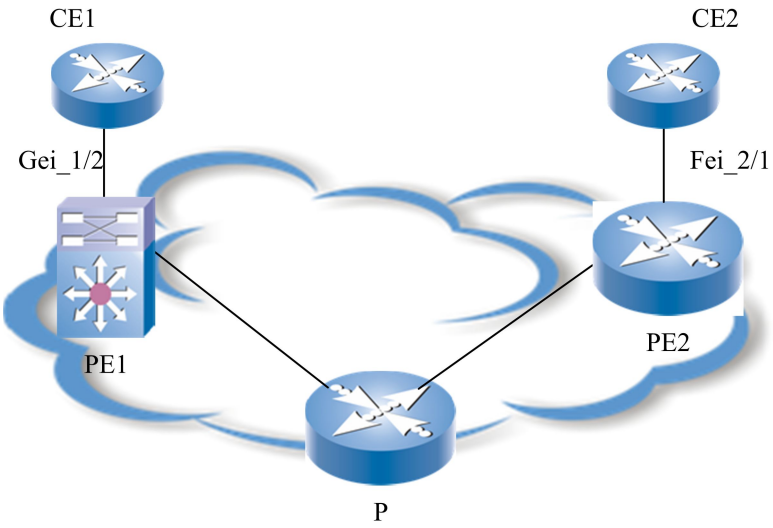
Note:

For step 2, if an IP address is configured on this interface in advance, after the VRF is associated with this interface, the IP address should be configured again.

Example This example shows how to configure VRF on PE.

As shown in Figure 20, R1, R2 and R3 are in the same MPLS domain. R1 is a device of ZXR10 8900 series switch. R2 and R3 are routers.

FIGURE 20 CONFIGURING VRF ON PE



Configuration on PE1:

```
PE1#configure terminal
PE1(config)#ip vrf vpn_a
PE1(config-vrf)#rd 100:1
PE1(config-vrf)#route-target import 100:1
PE1(config-vrf)#route-target export 100:1
PE1(config-vrf)#exit
PE1(config)#vlan 20
PE1(config-vlan)#switchport pvid gei_1/2
PE1(config-vlan)#exit
PE1(config)#interface vlan 20
PE1(config-if)#ip vrf forwarding vpn_a
PE1(config-if)#ip address 10.1.0.2 255.255.255.252
```

Configuration on PE2:

```
PE2#configure terminal
PE2(config)#ip vrf vpn_a
PE2(config-vrf)#rd 100:1
PE2(config-vrf)#route-target import 100:1
PE2(config-vrf)#route-target export 100:1
PE2(config-vrf)#exit
PE2(config)#interface fei_2/1
PE2(config-if)#ip vrf forwarding vpn_a
PE2(config-if)#ip address 10.2.0.2 255.255.255.252
```

Configuring Static Route between CE and PE

To configure a static route to CE on a PE, use the following command.

Command	Function
ZXR10(config)#**ip route vrf** *<vrf-name><prefix of destination ip address><network mask><next hop address>*	This configures a static route to CE on a PE

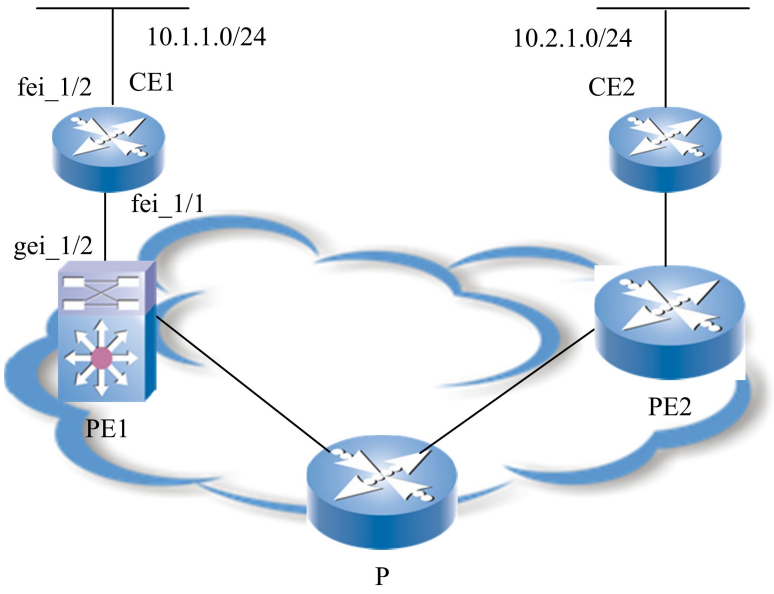
To redistribute static route to BGP, perform the following steps.

Step	Command	Function
1	ZXR10(config)#**router bgp** *<as-number>*	This enters BGP route configuration mode
2	ZXR10(config-router)#**address-family ipv4 vrf** *<vrf-name>*	This enters VRF address family configuration mode
3	ZXR10(config-router-af)#**redistribute static**	This redistributes the static route

Example This example shows how to configure static route between CE and PE.

As shown in Figure 21, configure a static on PE1 and CE1.

FIGURE 21 CONFIGURING STATIC ROUTE BETWEEN CE AND PE



Configuration on CE1:

```
CE1(config)#interface fei_1/1
CE1(config-if)#ip address 10.1.0.1 255.255.255.252
CE1(config)#interface fei_1/2
CE1(config-if)#ip address 10.1.1.254 255.255.255.0
CE1(config)#ip route 10.2.0.0 255.255.0.0 10.1.0.2
```

Configuration on PE1:

```
PE1(config)#ip route vrf vpn_a 10.1.0.0 255.255.0.0 10.1.0.1
PE1(config)#router bgp 100
PE1(config-router)#address-family ipv4 vrf vpn_a
PE1(config-router-af)#redistribute static
```

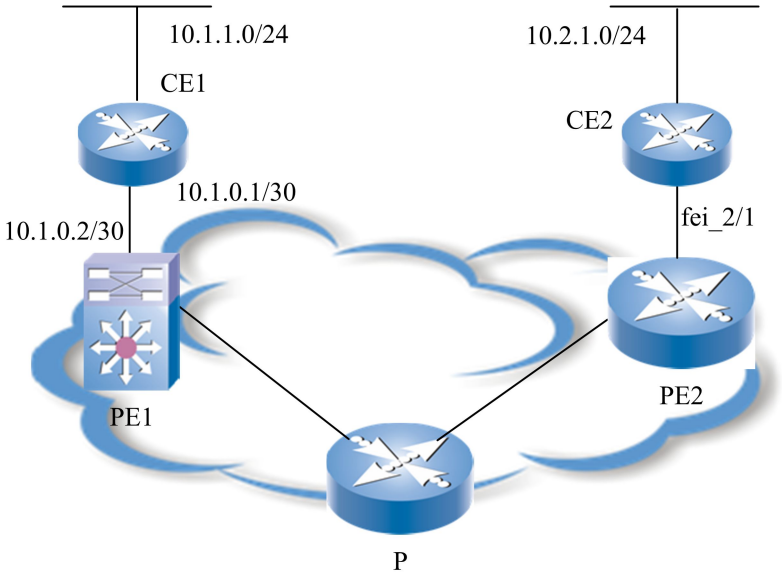
Configuring RIP between CE and PE

To configure RIP between CE and PE, perform the following steps.

Step	Command	Function
1	ZXR10(config)#**router rip**	This starts RIP process
2	ZXR10(config-router)#**version 2**	This configures RIP Version 2
3	ZXR10(config-router)#**address-family ipv4 vrf** *<vrf-name>*	This enters IPv4 address family configuration mode
4	ZXR10(config-router-af)#**no auto-summary**	This disables automatic aggregation function
5	ZXR10(config-router-af)#**version 2**	This configures RIP Version 2 in IPv4 address family configuration mode
6	ZXR10(config-router-af)#**network** *<network-number><wild-card>*	This enables RIP on an interface
7	ZXR10(config-router-af)#**redistribute connected**	This redistributes a direct route to RIP
8	ZXR10(config-router-af)#**redistribute bgp-int**	This redistributes IBGP to RIP

Example This example shows how to configure RIP between CE1 and PE1, as shown in Figure 22.

FIGURE 22 CONFIGURING RIP BETWEEN CE AND PE



Configuration on CE1:

```
CE1(config)#router rip
CE1(config-router)#no auto-summary
CE1(config-router)#version 2
CE1(config-router)#network  10.1.0.0  0.0.0.3
CE1(config-router)#redistribute connected
```

Configuration on PE1:

```
PE1(config)#router rip
PE1(config-router)#version 2
PE1(config-router)#address-family ipv4 vrf vpn_a
PE1(config-router-af)#no auto-summary
PE1(config-router-af)#version 2
PE1(config-router-af)#network 10.1.0.0 0.0.0.3
PE1(config-router-af)#redistribute bgp-int
PE1(config-router-af)#exit
PE1(config-router)#exit
PE1(config)#router bgp 100
PE1(config-router)#address-family ipv4 vrf vpn_a
PE1(config-router-af)#redistribute rip
PE1(config-router-af)#redistribute connected
```

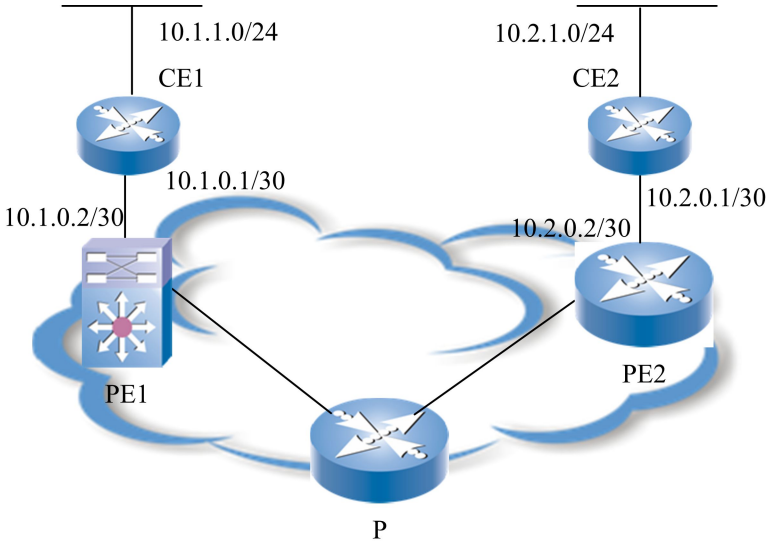
Configuring OSPF between CE and PE

To configure OSPF between a CE and a PE, perform the following steps.

Step	Command	Function
1	ZXR10(config)#**router ospf** <*process-id*> **vrf** <*vrf-name*>	This starts OSPF on a PE
2	ZXR10(config-router)#**network** <*network-number*><*wild-card*> **area** <*area-id*>	This enables OSPF on an interface
3	ZXR10(config-router)#**redistribute bgp-int**	This redistributes IBGP to OSPF

Example This example shows how to configure OSPF between CE1 and PE1, as shown in Figure 23.

FIGURE 23 CONFIGURING OSPF BETWEEN CE AND PE



Configuration on CE1:

```
CE1(config)#router ospf 1
CE1(config-router)#network 10.1.0.0 0.0.0.3 area 0.0.0.0
CE1(config-router)#network 10.1.1.0 0.0.0.255 area 0.0.0.0
```

Configuration on PE1:

```
PE1(config)#router ospf 2 vrf vpn_a
PE1(config-router)#network 10.1.0.0 0.0.0.3 area 0.0.0.0
PE1(config-router)#redistribute bgp-int
PE1(config-router)#exit
PE1(config)#router bgp 100
PE1(config-router)#address-family ipv4 vrf vpn_a
PE1(config-router-af)#redistribute ospf-int
PE1(config-router-af)#redistribute connected
```

Configurations on CE2 and PE2 are similar to that on CE1 and PE1.

Route information on CE1 is shown below.

```
CE1#show ip route ospf
IPv4 Routing Table:
Dest       Mask             Gw     Interface Owner  pri  metric
10.2.0.0 255.255.255.252 10.1.0.2  fei_1/1   ospf   110  1
10.2.1.0 255.255.255.0   10.1.0.2  fei_1/2   ospf   110  1
```

The result shows that CE1 has learned route from CE2 through OSPF.

Route information on CE2 is shown below.

```
CE2#show ip route ospf
IPv4 Routing Table:
Dest      Mask              Gw      Interface Owner  pri  metric
10.1.0.0 255.255.255.252 10.2.0.2   fei_1/1   ospf   110  1
10.1.1.0 255.255.255.0   10.2.0.2   fei_1/2   ospf   110  1
```

The result shows that CE2 has learned route from CE1 through OSPF.

CE1 can ping to the LAN connecting to CE2 successfully, as shown below.

```
CE1#ping 10.2.1.1
sending 5,100-byte ICMP echos to 10.2.1.1,timeout is 2 seconds.
!!!!!
Success rate is 100 percent(5/5),round-trip min/avg/max= 0/0/10
```

CE2 can ping to the LAN connecting to CE1 successfully, as shown below.

```
CE2#ping 10.1.1.1
sending 5,100-byte ICMP echos to 10.1.1.1,timeout is 2 seconds.
!!!!!
Success rate is 100 percent(5/5),round-trip min/avg/max= 0/0/10
```

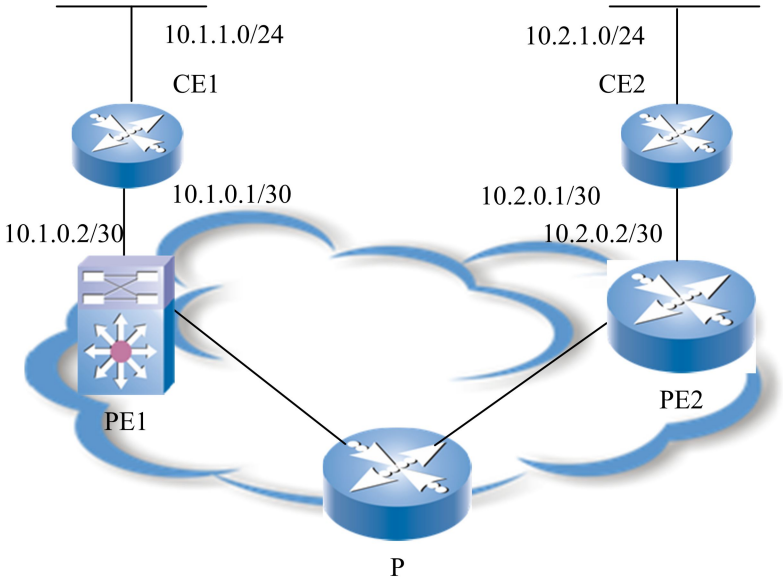
Configuring EBGP between CE and PE

To configure EBGP between a CE and a PE, perform the following steps.

Step	Command	Function
1	ZXR10(config)#**router bgp** <*as-number*>	This enters BGP route configuration mode
2	ZXR10(config-router)#**address-family ipv4 vrf** <*vpn-name*>	This enters IPv4 address family configuration mode
3	ZXR10(config-router-af)#**neighbor** <*ip-address*> **remote-as** <*as-number*>	This configures an EBGP peer
4	ZXR10(config-router-af)#**redistribute connected**	This redistributes direct route to BGP
5	ZXR10(config-router-af)#**exit-address-family**	This exits IPv4 address family configuration mode

Example This example shows how to configure EBGP between CE1 and PE1, as shown in Figure 24.

FIGURE 24 CONFIGURING EBGP BETWEEN CE AND PE



Configuration on CE1:

```
CE1(config)#router bgp 65001
CE1(config-router)#redistribute connected
CE1(config-router)#neighbor 10.1.0.2 remote-as 100
CE1(config-router)#neighbor 10.1.0.2 activate
```

Configuration on PE1:

```
PE1(config)#router bgp 100
PE1(config-router)#address-family ipv4 vrf vpn_a
PE1(config-router-af)#neighbor 10.1.0.1 remote-as 65001
PE1(config-router-af)#neighbor 10.1.0.1 activate
PE1(config-router-af)#redistribute connected
```

Configuring MPBGP

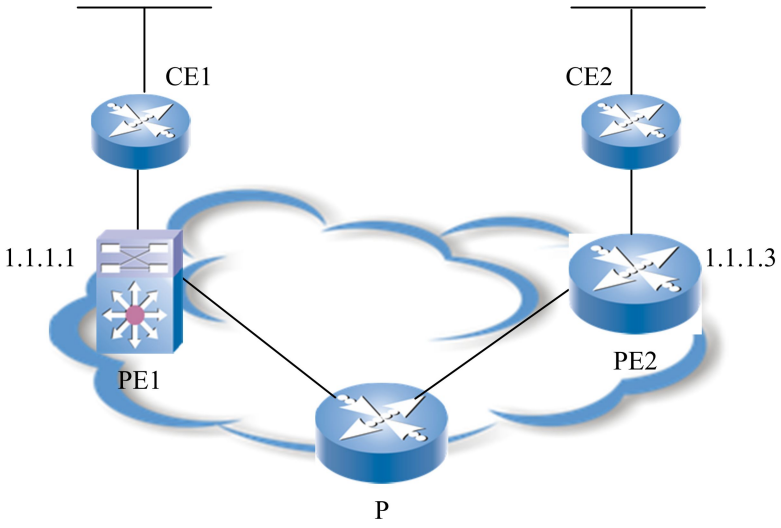
To configure MPBGP, perform the following steps.

Step	Command	Function
1	ZXR10(config)#**router bgp** *<as-number>*	This enters BGP route configuration mode
2	ZXR10(config-router)#**neighbor** *<ip-address>* **remote-as** *<as-number>*	This configures a BGP neighbor
3	ZXR10(config-router)#**neighbor** *<ip-address>* **activate**	This activates a neighbor
4	ZXR10(config-router)#**neighbor** *<ip-address>* **update-source loopback***<1-64>*	This configures the source address of updated route

Step	Command	Function
5	ZXR10(config-router)#**address-family ipv4 vrf** *<vpn-name>*	This enters IPv4 address family configuration mode
6	ZXR10(config-router-af)#**neighbor** *<ip-address>* **activate**	This activates a neighbor

Example This example shows how to configure MPBGP between PE1 and PE2, as shown in Figure 25.

FIGURE 25 CONFIGURING MPBGP



Configuration on PE1:

```
PE1(config)#router bgp 100
PE1(config-router)#neighbor 1.1.1.3 remote-as 100
PE1(config-router)#neighbor 1.1.1.3 activate
PE1(config-router)#neighbor 1.1.1.3 update-source loopback1
PE1(config-router)#address-family vpnv4
PE1(config-router-af)#neighbor 1.1.1.3 activate
```

Configuration on PE2:

```
PE2(config)#router bgp 100
PE2(config-router)#neighbor 1.1.1.1 remote-as 100
PE2(config-router)#neighbor 1.1.1.1 activate
PE2(config-router)#neighbor 1.1.1.1 update-source loopback1
PE2(config-router)#address-family vpnv4
PE2(config-router-af)#neighbor 1.1.1.1 activate
```

Configuring MPLS L3VPN Load Balance

To configure MPLS L3VPN load balance, perform the following steps.

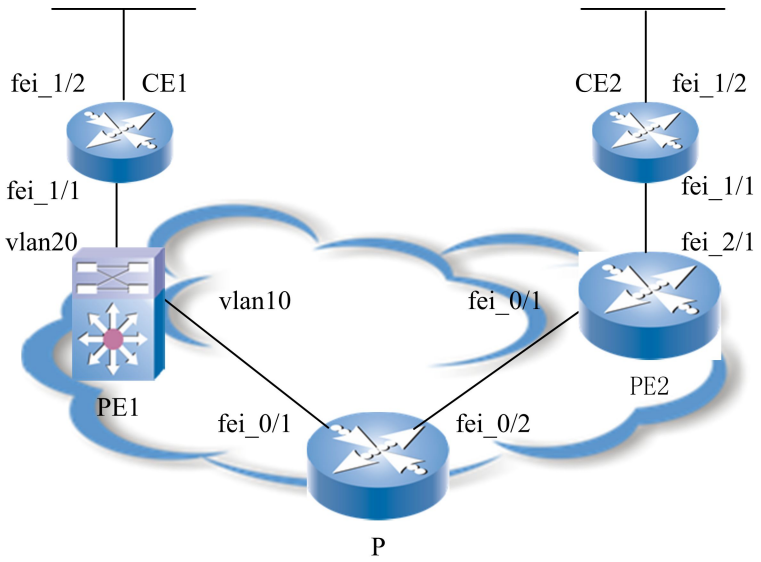
Step	Command	Function
1	`ZXR10(config)#`**router bgp** *<number>*	This enters BGP configuration mode
2	`ZXR10(config-router)#`**address-family ipv4 vrf**	This enters address VPN IPv4 configuration mode
3	`ZXR10(config-router-af)#`**vpn maximum-paths** *<number>*	This enables load balance function and sets the number of balanced routes

To disable load balance function, use **no vpn maximum-paths** command.

MPLS L3VPN Configuration Example

As shown in Figure 26, in a MPLS domain, PE1, P and PE2 run MPBGP, MPLS and OSPF. CE1 and PE1 run BGP. CE2 and PE2 run OSPF.

FIGURE 26 MPLS L3VPN CONFIGURATION EXAMPLE



Configuration on CE1:

```
CE1(config)#interface fei_1/1
CE1(config-if)#ip address 10.1.0.1 255.255.255.252
CE1(config-if)#exit
CE1(config)#interface fei_1/2
CE1(config-if)#ip address 10.1.1.254 255.255.255.0
CE1(config-if)#exit
CE1(config)#router bgp 65001
CE1(config-router)#redistribute connected
CE1(config-router)#neighbor 10.1.0.2 remote-as 100
CE1(config-router)#neighbor 10.1.0.2 activate
```

Configuration on PE1:

```
PE1(config)#ip vrf vpn_a
PE1(config-vrf)#rd 100:1
PE1(config-vrf)#route-target import 100:1
PE1(config-vrf)#route-target export 100:1
PE1(config)#vlan 10
PE1(config-vlan)#switchport pvid gei_1/1
PE1(config-vlan)#exit
PE1(config)#vlan 20
PE1(config-vlan)#switchport pvid gei_1/2
PE1(config-vlan)#exit
PE1(config)#interface vlan 10
PE1(config-if)#ip address 175.0.0.1 255.255.255.252
PE1(config-if)#mpls ip
PE1(config-if)#exit
PE1(config)#mpls ip
PE1(config)#interface vlan 20
PE1(config-if)#ip vrf forwarding vpn_a
PE1(config-if)#ip address 10.1.0.2 255.255.255.252
PE1(config-if)#exit
PE1(config)#router bgp 100
PE1(config-router)#neighbor 1.1.1.3 remote-as 100
PE1(config-router)#neighbor 1.1.1.3 activate
PE1(config-router)#neighbor 1.1.1.3 update-source loopback1
PE1(config-router)#address-family ipv4 vrf vpn_a
PE1(config-router-af)#redistribute connected
PE1(config-router-af)#neighbor 10.1.0.1 remote-as 65001
PE1(config-router-af)#neighbor 10.1.0.1 activate
PE1(config-router-af)#exit
PE1(config-router)#address-family vpnv4
PE1(config-router-af)#neighbor 1.1.1.3 activate
PE1(config-router-af)#exit
PE1(config-router)#exit
PE1(config)#router ospf 1
PE1(config-router)#network 1.1.1.1 0.0.0.0 area 0
PE1(config-router)#network 175.0.0.0 0.0.0.3 area 0
```

Configuration on P:

```
P(config)#interface fei_0/1
P(config-if)#ip address 175.0.0.2 255.255.255.252
P(config-if)#mpls ip
P(config-if)#exit
P(config)#interface fei_0/2
P(config-if)#ip address 176.0.0.1 255.255.255.252
P(config-if)#mpls ip
P(config-if)#exit
P(config)#mpls ip
P(config)#router ospf 1
P(config-router)#network 175.0.0.0 0.0.0.3 area 0
P(config-router)#network 176.0.0.0 0.0.0.3 area 0
```

Configuration on PE2:

```
PE2(config)#ip vrf vpn_a
PE2(config-vrf)#rd 100:1
PE2(config-vrf)#route-target import 100:1
PE2(config-vrf)#route-target export 100:1
PE2(config-vrf)#exit
```

```
PE2(config)#interface loopback1
PE2(config-if)#ip address 1.1.1.3 255.255.255.255
PE2(config-if)#exit
PE2(config)#interface fei_2/1
PE2(config-if)#ip vrf forwarding vpn_a
PE2(config-if)#ip address 10.2.0.2 255.255.255.252
PE2(config-if)#exit
PE2(config)#interface fei_0/1
PE2(config-if)#ip address 176.0.0.2 255.255.255.252
PE2(config-if)#mpls ip
PE2(config-if)#exit
PE2(config)#mpls ip
PE2(config)#router bgp 100
PE2(config-router)#neighbor 1.1.1.1 remote-as 100
PE2(config-router)#neighbor 1.1.1.1 activate
PE2(config-router)#neighbor 1.1.1.1 update-source loopback1
PE2(config-router)#address-family ipv4 vrf vpn_a
PE2(config-router-af)#redistribute ospf-int
PE2(config-router-af)#redistribute connected
PE2(config-router-af)#exit
PE2(config-router)#address-family vpnv4
PE2(config-router-af)#neighbor 1.1.1.1 activate
PE2(config-router-af)#exit
PE2(config-router)#exit
PE2(config)#router ospf 1
PE2(config-router)#network 1.1.1.3 0.0.0.0 area 0
PE2(config-router)#network 176.0.0.0 0.0.0.3 area 0
PE2(config-router)#exit
PE2(config)#router ospf 2 vrf vpn_a
PE2(config-router)#network 10.2.0.0 0.0.0.3 area 0
PE2(config-router)#redistribute bgp-int
```

Configuration on CE2:

```
CE2(config)#interface fei_1/1
CE2(config-if)#ip address 10.2.0.1 255.255.255.252
CE2(config-if)#exit
CE2(config)#interface fei_1/2
CE2(config-if)#ip address 10.2.1.254 255.255.255.0
CE2(config-if)#exit
CE2(config)#router ospf 1
CE2(config-router)#network 10.2.0.0 0.0.0.3 area 0
CE2(config-router)#network 10.2.1.0 0.0.0.255 area 0
```

CE1 can ping to the LAN connecting to CE2 successfully, as shown below.

```
CE1#ping 10.2.1.254
sending 5,100-byte ICMP echos to 10.2.1.254,timeout is 2 seconds.
!!!!!
Success rate is 100 percent(5/5),round-trip min/avg/max= 0/0/0
```

PE1 can ping to the LAN connecting to CE1 successfully, as shown below.

```
PE1#ping vrf vpn_a 10.1.1.254
sending 5,100-byte ICMP echos to 10.1.1.254,timeout is 2 seconds.
!!!!!
Success rate is 100 percent(5/5),round-trip min/avg/max= 0/0/0 ms.
```

PE1 can telnet to PE1, as shown below.

```
PE1#telnet 10.1.0.1 vrf vpn_a
```

VRF table on PE1 is shown below.

```
PE1#show ip route vrf vpn_a
IPv4 Routing Table:
Dest       Mask              Gw    Interface Owner   pri  metric
10.1.0.0  255.255.255.252  10.1.0.2  vlan20   direct   0    0
10.1.0.2  255.255.255.255  10.1.0.2  vlan20   address  0    0
10.1.1.0  255.255.255.0    10.1.0.1  vlan20   bgp      20   0
```

```
10.2.0.0  255.255.255.252  1.1.1.3   vlan10   bgp      200  0
10.2.1.0  255.255.255.0    1.1.1.3   vlan10   bgp      200  0
```

VRF table on PE2 is shown below.

```
PE2#show ip route vrf vpn_a
IPv4 Routing Table:
Dest       Mask              Gw    Interface  Owner  pri  metric
10.1.0.0  255.255.255.252  1.1.1.1   fei_0/1    bgp   200   0
10.1.1.0  255.255.255.0    1.1.1.1   fei_0/1    bgp   200   0
10.2.0.0  255.255.255.252  10.2.0.2  fei_2/1   direct  0    0
10.2.0.2  255.255.255.255  10.2.0.2  fei_2/1   address 0    0
10.2.1.0  255.255.255.0    10.2.0.1  fei_2/1    ospf   110  2
```

EBGP connection between CE1 and PE1 is shown below.

```
CE1#show ip bgp summary
Neighbor  Ver  As  MsgRcvd  MsgSend  Up/Down(s)  State/PfxRcd
10.1.0.2   4   100  235       234       02:00:22      3
```

OSPF information between CE2 and PE2 is shown below.

```
PE2#show ip ospf neighbor process 2
          OSPF Router with ID (10.2.0.2) (Process ID 2)
Neighbor 10.2.0.1
    In the area 0.0.0.0
    via interface fei_2/1 10.2.0.1
    Neighbor is BDR
    State FULL, priority 1, Cost 1
    Queue count : Retransmit 0, DD 0, LS Req 0
    Dead time : 00:00:29 Options : 0x42
    In Full State for 02:02:42
```

IBGP connection between PE1 and PE2 is shown below.

```
PE1#show ip bgp summary
Neighbor  Ver  As    MsgRcvd  MsgSend  Up/Down(s)  State/PfxRcd
1.1.1.3    4   100    139       132       01:09:33      0
10.1.0.1   4   65001  240       241       02:03:36      2
```

The Intags distributed by PE1 for VRF route are shown below.

```
PE1#show ip protocol routing vrf vpn_a
Routes of vpn:
status codes: *valid, >best, s-stale
      Dest        NextHop   Intag   Outtag   RtPrf    Protocol
*>  10.1.0.0/30  10.1.0.2    17     notag     0       connected
*>  10.1.0.2/32  10.1.0.2    16     notag     0       connected
*>  10.1.1.0/24  10.1.0.1    19     notag     20       bgp-ext
*>  10.2.0.0/30  1.1.1.3     22     17       200       bgp-int
*>  10.2.1.0/24  1.1.1.3     23     19       200       bgp-int
```

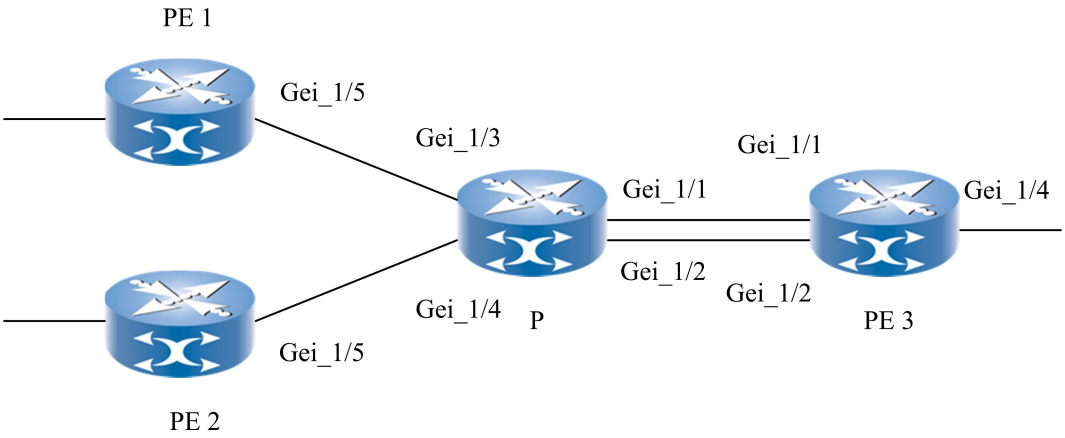
The Intags of VRF routes on PE2 are shown below.

```
PE2#sh ip protocol routing vrf vpn_a
Routes of vpn:
status codes: *valid, >best, s-stale
     Dest        NextHop   Intag    Outtag   RtPrf    Protocol
*>  10.1.0.0/30  1.1.1.1    22       17      200      bgp-int
*>  10.1.1.0/24  1.1.1.1    23       19      200      bgp-int
*>  10.2.0.0/30  10.2.0.2   17      notag    0        connected
*   10.2.0.0/30  10.2.0.0   18      notag    110      ospf
*>  10.2.0.2/32  10.2.0.2   16      notag    0        connected
*>  10.2.1.0/24  10.2.0.1   19      notag    110      ospf
```

Configuration Example of MPLS L3VPN Load Balance on Smartgroup Link

As shown in Figure 27, PE1 and PE2 bear multiple private VLANs. PE1 and PE2 connect to P. There are two links between P and PE3. PE1, PE2 and PE3 runs LDP.

FIGURE 27 CONFIGURATION EXAMPLE OF MPLS L3VPN LOAD BALANCE ON SMARTGROUP LINK



Configuration points:

- Configure interface smartgroup1 on P, including gei_1/1 and gei_1/2. Configure interface smartgroup1 on PE3, including gei_1/1 and gei_1/2.
- Enable VPN load balance function on smartgroup1 of P and PE3. Configure to Implement load balance mode according to private network VLAN.
- View the configuration on P and PE3 with **show** command.

PE3 configuration:

```
PE3(config)#interface smartgroup10
PE3(config-if)#switchport mode trunk
PE3(config-if)#exit
PE3(config)#interface gei_1/1
PE3(config-if)#switchport mode trunk
PE3(config-if)#switchport trunk vlan 1
PE3(config-if)#smartgroup 10 mode active
PE3(config-if)#exit
PE3(config)#interface gei_1/2
PE3(config-if)#switchport mode trunk
PE3(config-if)#switchport trunk vlan 1
PE3(config-if)#smartgroup 10 mode active
PE3(config-if)#exit
PE3(config)#interface smartgroup10
PE3(config-if)#smartgroup load-balance pri-label
PE3(config-if)#show running-config interface smartgroup10
Building configuration...
!
interface smartgroup10
  out_index 138
  switchport mode trunk
  switchport trunk native vlan 1
```

```
  smartgroup load-balance pri-label
  smartgroup mode 802.3ad
!
end
```

Configuration on P:

```
P(config)#interface smartgroup10
P(config-if)#switchport mode trunk
P(config-if)#exit
P(config)#interface gei_1/1
P(config-if)#switchport mode trunk
P(config-if)#switchport trunk vlan 1
P(config-if)#smartgroup 10 mode active
P(config-if)#exit
P(config)#interface gei_1/2
P(config-if)#switchport mode trunk
P(config-if)#switchport trunk vlan 1
P(config-if)#smartgroup 10 mode active
P(config-if)#exit
P(config)#interface smartgroup10
P(config-if)#smartgroup load-balance pri-label
P(config-if)#show running-config interface smartgroup10
Building configuration...
!
interface smartgroup10
  out_index 138
  switchport mode trunk
  switchport trunk native vlan 1
smartgroup load-balance pri-label
  smartgroup mode 802.3ad
!
end
```

Chapter 4

MPLS L2VPN Configuration

Table of Contents

MPLS L2VPN Overview 51
Configuring MPLS L2VPN 54
MPLS L2VPN Configuration Examples.....57
MPLS L2VPN Maintenance and Diagnosis.....62

MPLS L2VPN Overview

L2VPN is a solution for transporting Layer 2 traffic over an IP/MPLS backbone. It extends the usability of the IP/MPLS backbone by enabling it to offer both Layer 2 and Layer 3 services. It can accommodate many types of Layer 2 frames, including Ethernet, Frame Relay, ATM, PPP and HDLC.

Service provider provides two forms of L2VPN services:

- VPWS
- VPLS

Virtual Private Wire Service

Virtual Private Wire Service (VPWS) is based on infrastructure of MPLS net, providing high-speed Layer 2 transparent transmission between a pair of ports of the two routers. VPWS is mainly composed of PE router, LDP and LSP Tunnel of the MPLS.

PE router possesses and maintains link information of Layer 2 transparent transmission connected directly to it. PE router is responsible for making and removing labels on common packet of VPN clients, so that PE router should be an edge label switch router.

Two ports of Layer 2 transparent transmission between two PE routers are of same type like Ethernet, VLAN, and ATMVC, frame-relay VC, HDLC or PPP. Each pair of such ports are represented by sole VC Label VCID.

LSP tunnel through MPLS net should be defined between two PE routers and should provide Tunnel Label transparently transmitting

data between two PE routers. At the same time, direct process of LDP label distribution protocol is also defined between two PE routers to transmit virtual link information. Among them, distributing VC Label through matching VCID is critical.

When data packet enters PE router at the port of Layer 2 transparent transmission, PE router finds the corresponding Tunnel Label and VC Label through matching VCID. PE router will put two layers labels on the data packet. External layer is Tunnel Label indicating the route from this PE router to destination PE router. Internal layer is VC Label indicating which corresponding router port of VCID belongs to on destination PE router.

PE router should monitor Layer 2 protocol state at each port, such as frame-relay LMI and ILMI of ATM. When a fault occurs, users can cancel VC Label through LDP label distribution protocol process so that Layer 2 transparent transmission is shut off avoiding producing unidirectional unwanted data stream.

Such Layer 2 transparent transmission based on MPLS changes traditional confinement that Layer 2 link should be implemented through network switch. It essentially forms a pattern of One Net Multi-Service pattern and makes the operator provide Layer 2 and Layer 3 Services simultaneously in a MPLS net.

Virtual Private LAN Service

Virtual Private LAN Service (VPLS) allows service providers to deliver VPN services based on Ethernet with the same level of support and reliability as existing services such as Frame Relay and ATM. VPLS combines the cost-effectiveness of Ethernet with the service delivery, traffic engineering, scalability and reliability of Multi Protocol Label Switching (MPLS) to provide a multipoint Layer 2 Ethernet VPN service.

VPNs are a value-added premium service that a service provider can offer to its business customers. VPNs allow service providers to connect multiple customer sites together securely over a common shared network. The service keeps individual customer data private from other customers' data using virtual connections across the shared network. Traditionally, VPNs are based on leased lines, Frame Relay, or ATM services, but more recently on Ethernet, IP and Multi Protocol Label Switching (MPLS).

VPNs are major source of revenue for service providers as they provide flexibility, scalability and integrate well with data services. However, VPNs use traditional bandwidth delivery services to connect to the customer, and are costly to provision and complex to manage. The continued growth of the Internet and IP-based services combined with a move by customers to outsource more IT functions, such as storage and backup, is placing a strain on these fixed-bandwidth VPN delivery models.

Ethernet provides simple, flexible, more scalable bandwidth than traditional fixed bandwidth solutions and has revolutionized Metro environment. For the end-user, Ethernet is less expensive, easier to connect and simpler to manage. For service provider, it is cheaper, more flexible, and allows new services to be provisioned more simply and quickly than traditional solutions.

Service providers which provide two-layer VPN of MPLS are only in charge of the connectivity of layer two, and they do not take part in routing computation of VPN subscription. All-connected two-layer VPN has a problem that it needs to assign a connection between CE and PE for every CE to others.

For example, if there are four CEs in one VPN, it needs to establish six connections for all CEs. The extension of two-layer VPN of MPLS is not good.

Two-layer VPN can be built through MP-BGP extension or LDP extension, and drafts of these two are respectively draft-ietf-l2vpn-vpls-bgp-xx and draft-ietf-l2vpn-vpls-ldp-xx.

In VPLS LDP plan, it emphasizes in solving the problem of how to establish Virtual Circuit (VC) between two CEs. Martini L2 VPN uses VC-TYPE+VC-ID to distinguish VC. VC-TYPE indicates type of VC is ATM, VLAN or PPP and VC-ID is used to identity a VC solitarily.

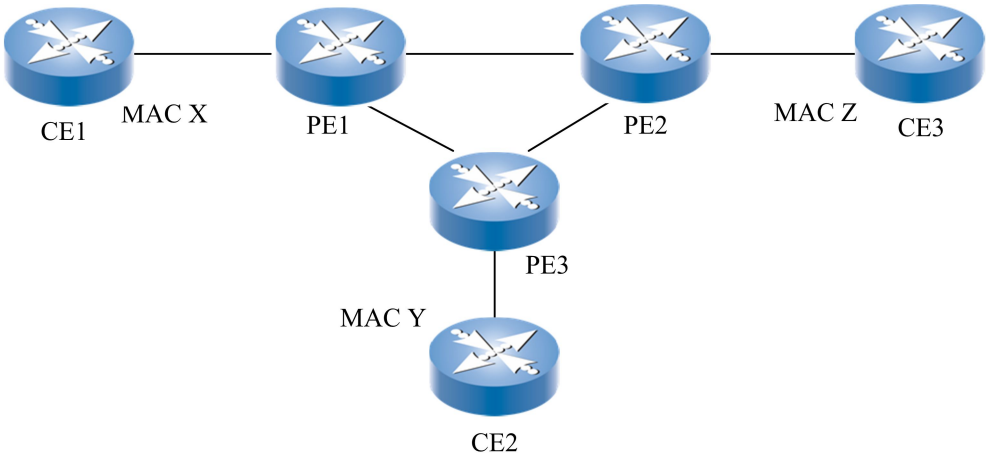
In all VCs of one VC-TYPE, VC-ID has uniqueness in the entire SP network. PE which connects to two CEs switches VC-ID through LDP and binds CE with its VC-ID.

After LSP has been established successfully between two PEs and VC-IDs of two sides have been switched and bound, a VC is established successfully. Two CEs can transmit two-layer data through this VC.

For switching VC-ID between PEs, Martini draft extended LDP and added FEC type (128 types and 129 types) of VC FEC. Moreover, two PEs which switches VC-ID may be not connected directly, LDP must use Remote peer to establish session and switch VC FEC and VC-ID in this session.

CE1, CE2 and CE3 are connected through core network of MPLS. In the view of subscribers, they are connected through a LAN network. This is shown in Figure 28.

FIGURE 28 NETWORK OF VPLS



VPLS instances of PE1, PE2 and PE3 are connected to others and build an all-connected of PW. All VPLS instances in one VPLS domain use the same VC-ID.

It is assumed that PE1 distributes a VC label of 102 to PE2 and a VC label of 103 to PE3. PE2 distributes a VC label of 201 to PE1

and a VC label of 203 to PE3. PE3 distributes a VC label of 301 to PE1 and a VC label of 302 to PE2.

A host which connects to CE1 sends a frame with source MAC address X and destination MAC address Y to PE1. PE1 does not know which PE owns MAC address Y. It adds label of 201 in this frame and sends it to PE2. Similarly, it adds the label of 301 in this frame and sends it to PE3.

After PE2 receives this frame, it learns source MAC address X of this frame and binds MAC address X with VC label of 102 which was distributed by PE1.

Configuring MPLS L2VPN

Configuring VPWS

To configure VPWS, perform the following steps:

Step	Command	Function
1	`ZXR10(config)#`**netcard vpn**	This configures micro code version. It is required to reboot line interfaces with network processor after this command is configured.
2	`ZXR10(config-if)#`**mpls xconnect** [**static-pw**]<*ip-address*>[<*static-p*>]<*vc-id*>[**tunnel** <*tunnel number*>][**mtu** <*mtu value*>]	This starts VPWS configuration on layer 3 VLAN interface.

Configure remote interface address and tunnel ID, establish VPWS between local interface and remote interface. Vc-id must be unique on local node and the corresponding interface on remote node must have the same vc-id. Network-unique PW identifier is composed of remote node IP address and vc-id. With parameter tunnel number, MPLS label is assigned by specified tunnel, without parameter tunnel, the label is assigned by LDP. Mtu value can be specified by mtu parameter.

Configuring VPLS

To configure VPLS, perform the following steps.

Step	Command	Function
1	ZXR10(config)#**netcard vpn**	This configures micro code version. It is required to reboot line interfaces with network processor after this command is configured. When this command is saved, reboot the rack instead of line interfaces.
2	ZXR10(config)#**mac learn special**	This configures MAC address study mode.
3	ZXR10(config)#**vfi** <*vfi-name*>	This creates VFI and enters VFI configuration mode.
4	ZXR10(config-vfi)#**vcid** <*vcid*>	This configures vcid. when VPLS instances of two PE routers belong to same VPLS region, it is required to establish PW to continue the communication between VPLS instances. PE establishes PW to identify mpls lining-label (or called VC label. Mpls outer-label, which compose lsp, together with VC label, compose the PW) of VPLS region with each other. Vc-id configured by vcid is used to identify the VPLS region when PE changes VC label.
5	ZXR10(config-vfi)#**pwtype** { **ethernet** | **ethernet-vlan**}	This configures pwtype parameter. Two PEs need to negotiate PW type when they establish PW. There are two types of PWs in VPLS which are ethernet and ethernet-vlan. When PW types are different with each other, PW can not be established.
6	ZXR10(config-vfi)#**peer** <*peer-ip-address*>[**spoke**]	This configures peer parameters. It is required to designate all PE addresses in one VPLS region. Spoke type is used to bed VPLS and span regions.
7	ZXR10(config-if)#**xconnect vfi** <*vfi-name*>	This starts VPLS configuration on L3 interface.
8	ZXR10(config)#**mpls ldp target-session** <*peer-ip-address*>	This configures expanding LDP neighbor. VPLS module function is based on LDP protocol. To establish PW between indirect PEs, it is important to establish LDP neighbor through switching TARGET HELLO of LDP, and then distribute label of PW.

Step	Command	Function
9	ZXR10#**clear mac-table vfi** <*vfi-name*>	This configures to clear the corresponding MAC address table of designated VPLS on privileged mode.
10	ZXR10(config)#**mac learning-strategy micode**	This configures MAC learning mode.
11	ZXR10(config)#**mac learn special**	This changes MAC learning mode to special.
12	ZXR10(config)#**static-pw** <*pw-name*><*local-label*><*remote-label*>	This configures static PW entry on local NE.
13	ZXR10(config-if)#**#mpls xconnect** <*peer-address*> **static-pw** <*pw-name*>	This configures L2VPN VPWS service using static PW. By default, TMPLS OAM function of MEG is not enabled.
14	ZXR10(config-vfi)#**peer static-pw** <*peer-ip-address*><*static-pw-name*>[**spoke**][**tunnel** <*tunnel-number*>]	This configures L2VPN VPLS service on VFI configuration mode. The default attribute is hub.
15	ZXR10(config-vfi)#**mtu** <*mtu-value*>	This configures mtu of VFI on VFI configuration mode.
16	ZXR10(config)#**show static pw** [<*pw-name*>| **brief**]]	This views static PW entry information.

Configuring Layer 2 Control Message Transparent Transmission VPLS Network Policy

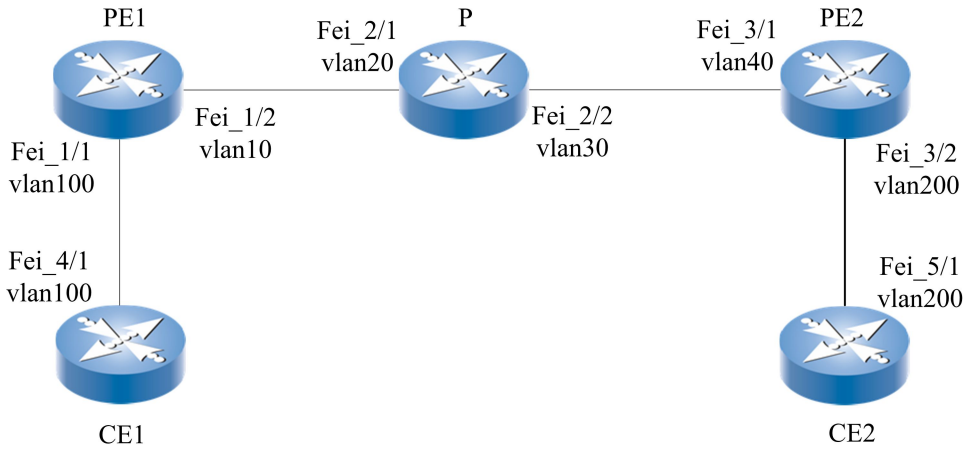
Step	Command	Function
1	ZXR10(config-vfi)#**vcid** < *vcid*>	This configures VFI parameter vcid.
2	ZXR10(config-vfi)#**pwtype** { **ethernet**| **ethernet-vlan**}	This configures VFI parameter pwtype.
3	ZXR10(config-vfi)#**l2protocol-tunnel** { **mac-control** | **lacp** | **port-auth** | **lldp** | **garp** | **stp** | **all-bridge-multicast** | **zdp-ztp** | **all**}	This configures layer 2 message transparent transmission VPLS network forwarding policy.
4	ZXR10(config-vfi)#**show vfi** <*vfi-name*>	This views a specific vfi detailed information.

MPLS L2VPN Configuration Examples

VPWS Configuration Example

A network of VPWS is shown in Figure 29.

FIGURE 29 VPWS CONFIGURATION EXAMPLE



The configuration requirements are as follows:

- Configure interface addresses on fei_1/2 of PE1, fei_2/1 and fei_2/2 of P, and fei_3/1 of PE2.
- Configure loopback addresses for PE1, P and PE2.
- Run IGP protocol (such as OSPF) on PE1, P and PE2 to implement interconnection of PE1 and PE2 and to learn the loopback interface address route of the peer party.
- Start MPLS on PE1, P and PE2 and specify router-id of mpls ldp. Start mpls ip on vlan10 of PE1, vlan20 and vlan30 of P, and vlan40 of PE2.
- Configure target—session on PE1 and PE2 to establish ldp neighborhood relationship between PE1 and PE2. If P does not exist in networking, the configuration is unnecessary.
- Configure mpls xconnect on interfaces vlan100 and vlan200 on PE1 and PE2 connecting CE.

The following describes the configurations of each device.

PE1 configuration:

```
PE1(config)#interface loopback10
PE1(config-if)#ip address  1.1.1.1 255.255.255.255
PE1(config)#interface vlan100
PE1(config-if)#mpls xconnect 1.1.1.3 100
PE1(config)#interface vlan10
PE1(config-if)#ip address  175.1.1.1 255.255.255.0
PE1(config-if)#mpls ip
PE1(config)#mpls ip
PE1(config)#mpls ldp router-id loopback10 force
```

```
PE1(config)#mpls ldp target-session 1.1.1.3
PE1(config)#router ospf 1
PE1(config-router)#network 1.1.1.1 0.0.0.0 area 0.0.0.0
PE1(config-router)#network 175.1.1.0 0.0.0.255 area 0.0.0.0
```

P configuration:

```
P(config)#interface loopback10
P(config-if)#ip address  1.1.1.2 255.255.255.255
P(config)#interface vlan20
P(config-if)#ip address  175.1.1.2 255.255.255.0
P(config-if)#mpls ip
P(config)#interface vlan30
P(config-if)#ip address  148.1.1.2 255.255.255.0
P(config-if)#mpls ip
P(config)#mpls ip
P(config)#mpls ldp router-id loopback10 force
P(config)#router ospf 1
P(config-router)#network 1.1.1.2 0.0.0.0 area 0.0.0.0
P(config-router)#network 148.1.1.0 0.0.0.255 area 0.0.0.0
P(config-router)#network 175.1.1.0 0.0.0.255 area 0.0.0.0
```

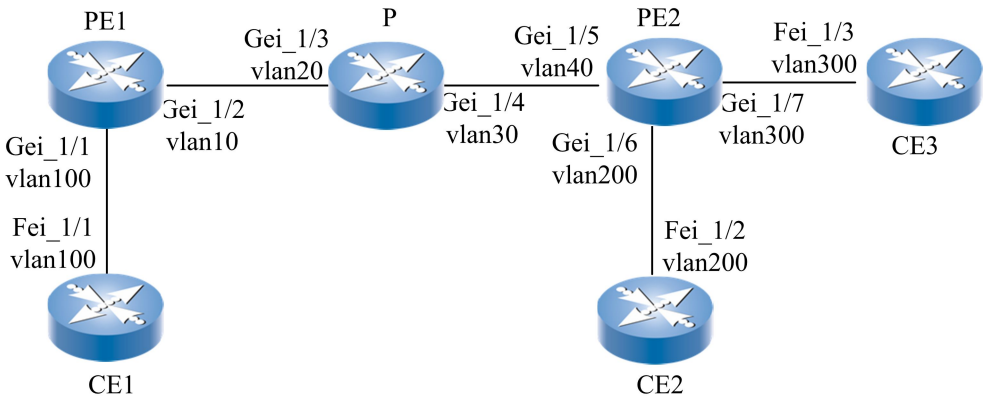
PE2 configuration:

```
PE2(config)#interface loopback10
PE2(config-if)#ip address  1.1.1.3 255.255.255.255
PE2(config)#interface vlan40
PE2(config-if)#ip address  148.1.1.3 255.255.255.0
PE2(config-if)#mpls ip
PE2(config)#interface vlan200
PE2(config-if)#mpls xconnect 1.1.1.1 100
PE2(config)#mpls ip
PE2(config)#mpls ldp router-id loopback10 force
PE2(config)#mpls ldp target-session 1.1.1.1
PE2(config)#router ospf 1
PE2(config-router)#network 1.1.1.3 0.0.0.0 area 0.0.0.0
PE2(config-router)#network 148.1.1.0 0.0.0.255 area 0.0.0.0
```

VPLS Configuration Example

VPLS L2VPN interconnection needs to be implemented among CE1, CE2 and CE3. This is shown in Figure 30.

FIGURE 30 VPLS CONFIGURATION EXAMPLE 1



The configuration is as follows:

- Create VFI and configure VFI parameters such as vcid, pwtype and peer.
- Bind physical address to VLAN, as shown in the figure, bind gei_1/1 of PE1 to vlan 10, and gei_1/2 to vlan20.
- Configure interface addresses on vlan10 of PE1, vlan20 and vlan30 of P, and vlan40 of PE2.
- Configure loopback addresses for PE1, P and PE2.
- Run IGP protocol (such as OSPF) on PE1, P and PE2 to implement interconnection of PE1 and PE2 and to learn loopback interface address route of the peer party.
- Start MPLS on PE1, P and PE2: Configure mpls ip globally and specify the router-id of mpls ldp. Start mpls ip on vlan10 of PE1, vlan20 and vlan30 of P, and vlan40 of PE2.
- Configure target—session on PE1 and PE2 to establish the Idp neighborhood relationship between PE1 and PE2. If P does not exist in networking, the configuration is unnecessary.
- Configure xconnect vfi on the interfaces of vlan100, vlan200 and vlan300 on PE1 and PE2 connecting CE.

The following describes the configurations of each device.

PE1 configuration:

```
PE1(config)#vfi vpls_a
PE1(config-vfi)#vcid 100
PE1(config-vfi)#pwtype ethernet-vlan
PE1(config-vfi)#peer 1.1.1.3

PE1(config)#interface gei_1/1
PE1(config-if)#switch access vlan 100
PE1(config-if)#exit
PE1(config)#interface gei_1/2
PE1(config-if)#switch access vlan 10
PE1(config-if)#exit
PE1(config)#interface loopback10
PE1(config-if)#ip address  1.1.1.1 255.255.255.255
PE1(config)#interface vlan 100
PE1(config-if)#xconnect vfi vpls_a
PE1(config-if)#exit
PE1(config)#interface vlan 10
PE1(config-if)#ip address  175.1.1.1 255.255.255.0
PE1(config-if)#mpls ip
PE1(config-if)#exit
PE1(config)#mpls ip
PE1(config)#mpls ldp router-id loopback10 force
PE1(config)#mpls ldp target-session 1.1.1.3
PE1(config)#router ospf 1
PE1(config-router)#network 1.1.1.1 0.0.0.0 area 0.0.0.0
PE1(config-router)#network 175.1.1.0 0.0.0.255 area 0.0.0.0
```

P configuration:

```
P(config)#interface gei_1/3
P(config-if)#switch access vlan 20
P(config-if)#exit
P(config)#interface gei_1/4
P(config-if)#switch access vlan 30
P(config-if)#exit
P(config)#interface loopback10
P(config-if)#ip address  1.1.1.2 255.255.255.255
P(config-if)#exit
P(config)#interface vlan 20
P(config-if)#ip address  175.1.1.2 255.255.255.0
P(config-if)#mpls ip
P(config-if)#exit
P(config)#interface vlan 30
```

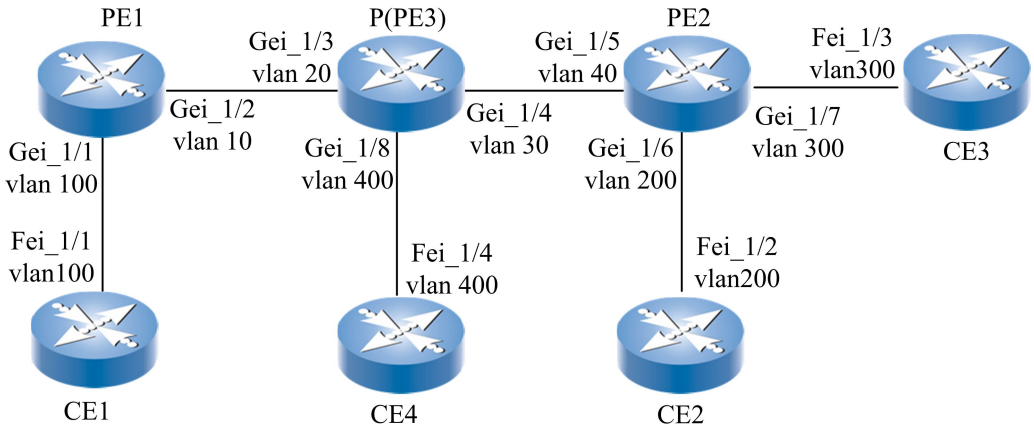
```
P(config-if)#ip address  148.1.1.2 255.255.255.0
P(config-if)#mpls ip
P(config-if)#exit
P(config)#mpls ip
P(config)#mpls ldp router-id loopback10 force
P(config)#router ospf 1
P(config-router)#network 1.1.1.2 0.0.0.0 area 0.0.0.0
P(config-router)#network 148.1.1.0 0.0.0.255 area 0.0.0.0
P(config-router)#network 175.1.1.0 0.0.0.255 area 0.0.0.0
```

PE2 configuration:

```
PE2(config)#vfi vpls_a
PE2(config-vfi)#vcid 100
PE2(config-vfi)#pwtype ethernet-vlan
PE2(config-vfi)#peer 1.1.1.1
PE2(config)#interface gei_1/5
PE2(config-if)#switch access vlan 40
PE2(config-if)#exit
PE2(config)#interface gei_1/6
PE2(config-if)#switch access vlan 200
PE2(config-if)#exit
PE2(config)#interface gei_1/7
PE2(config-if)#switch access vlan 300
PE2(config-if)#exit
PE2(config)#interface loopback10
PE2(config-if)#ip address  1.1.1.3 255.255.255.255
PE2(config)#interface vlan 40
PE2(config-if)#ip address  148.1.1.3 255.255.255.0
PE2(config-if)#mpls ip
PE2(config)#interface vlan 200
PE2(config-if)#xconnect vfi vpls_a
PE2(config-if)#exit
PE2(config)#interface vlan 300
PE2(config-if)#xconnect vfi vpls_a
PE2(config-if)#exit
PE2(config)#mpls ip
PE2(config)#mpls ldp router-id loopback10 force
PE2(config)#mpls ldp target-session 1.1.1.1
PE2(config)#router ospf 1
PE2(config-router)#network 1.1.1.3 0.0.0.0 area 0.0.0.0
PE2(config-router)#network 148.1.1.0 0.0.0.255 area 0.0.0.0
```

Suppose that networking in Figure 30 is changed. As shown in Figure 31, a CE4 is connected to P so that CE1, CE2, CE3 and CE4 are in the same VFI. In the networking, P can also be configured to a PE (PE3).

FIGURE 31 VPLS CONFIGURATION EXAMPLE 2



PE1 configuration:

```
PE1(config)#vfi vpls_a
PE1(config-vfi)#peer 1.1.1.2
```

P (PE3) configuration:

```
P(config)#vfi vpls_a
P(config-vfi)#vcid 100
P(config-vfi)#pwtype ethernet-vlan
P(config-vfi)#peer 1.1.1.1
P(config-vfi)#peer 1.1.1.3
P(config)#interface gei_1/8
P(config-if)#switch access vlan 400
P(config-if)#exit
P(config)#interface vlan 400
P(config-if)#xconnect vfi vpls_a
```

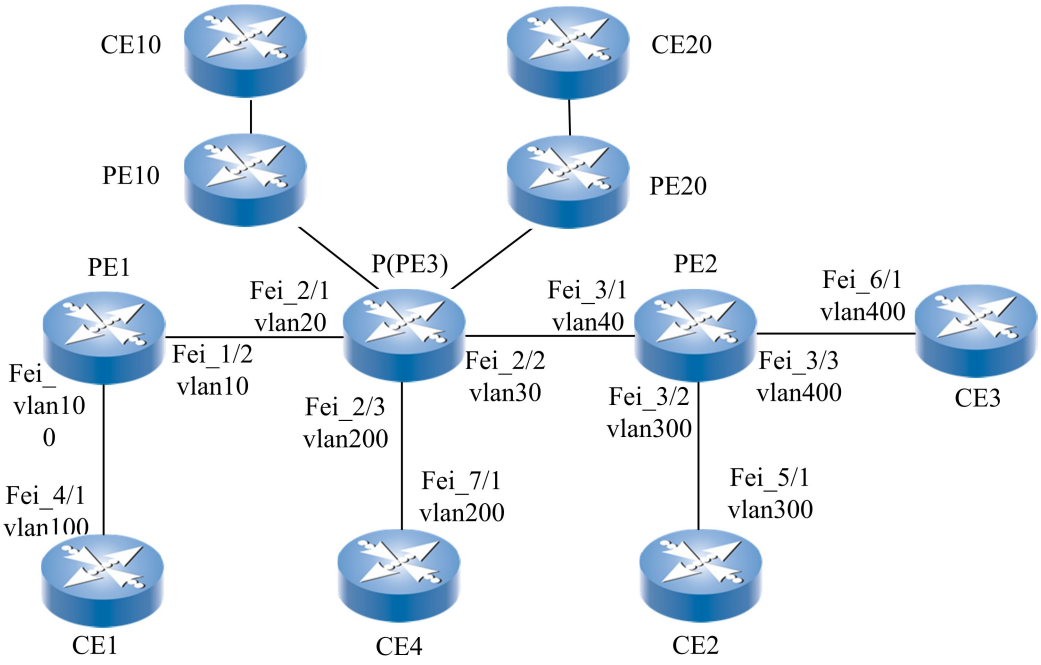
PE2 configuration:

```
PE2(config)#vfi vpls_a
PE2(config-vfi)#peer 1.1.1.2
```

In the two networking modes in Figure 30, and in Figure 31, the flat VPLS network is inside the MPLS network. PWs connecting PEs in a VPLS instance are HUB PWs.

When a new PE is added into the VPLS instance, no PW needs to be established between the PE and other PEs. As shown in Figure 32, PE10 and PE 20 only establish PW relationship with P (also called PE3). They do not need to establish PW relationship with PE1 and PE2. This PW type is called SPOKE PW.

FIGURE 32 VPLS CONFIGURATION EXAMPLE 3



Router IDs of PE10 and PE20 are 1.1.1.10 and 1.1.1.20. Configuration is as follows.

P (PE3) configuration:

```
P(config)#vfi vpls_a
P(config-vfi)#peer 1.1.1.10 spoke
P(config-vfi)#peer 1.1.1.20 spoke
```

For other configurations, refer to PE1.

PE10 configuration:

```
PE10(config)#vfi vpls_a
PE10(config-vfi)#peer 1.1.1.2 spoke
```

For other configurations, refer to PE1.

PE20 configuration:

```
PE20(config)#vfi vpls_a
PE20(config-vfi)#peer 1.1.1.2 spoke
```

Layer 2 Control Message Transparent Transmission VPLS Network Configuration Example

Layer 2 control message forwarding policy is applied in a vfi on a PE device, such as configuring lldp and stp message transparent transmission.

```
ZXR10(config)# vfi test
ZXR10(config-vfi)# vcid 1
ZXR10(config-vfi)# pwtype ethernet
ZXR10(config-vfi)# l2protocol-tunnel lldp
ZXR10(config-vfi)# l2protocol-tunnel stp
```

MPLS L2VPN Maintenance and Diagnosis

VPWS Maintenance and Diagnosis

To configure VPWS maintenance and diagnosis, perform the following steps.

Step	Command	Function
1	ZXR10#**show mpls l2transport vc** [{**vcid** <*vcidmin*>[<*vcidmax*>]| **interface** <*interface-name*>[<*loca-lcircuit-id1*>[<*local-circuit-id2*>]]| **destination** <*ip-addr*>}][**detail**]	This checks whether VC is established
2	ZXR10#**show mpls l2transport binding** [<*vc-id*>|<*ip-address*>|**local-label** < *local-label*>| **remote-label** <*remote-label* >]	This checks VC binding information

Step	Command	Function
3	ZXR10#**debug mpls ldp l2vpn event**	This monitors L2VPN event debug information
4	ZXR10#**debug mpls ldp l2vpn fsm**	This monitors state machine of L2VPN

VPLS Maintenance and Diagnosis

To configure VPLS maintenance and diagnosis, perform the following steps.

Step	Command	Function
1	ZXR10#**show vfi** <*vfi-name*>	This checks VFI information
2	ZXR10#**show mpls l2transport vc vpls** [{[**vcid** <*vc-id*>[<*max-vc-id*>]][**vfi** <*vfi-name*>][**destination** <*ip-address*>][**detail**]}]	This checks whether VC is established
3	ZXR10#**show mpls l2transport binding**	This checks VC binding information
4	ZXR10#**debug mpls ldp l2vpn event**	This monitors L2VPN event debug information
5	ZXR10#**debug mpls ldp l2vpn fsm**	This monitors state machine of L2VPN

Layer 2 Control Message Transparent Transmission VPLS Network Configuration Maintenance and Diagnosis

To maintain layer 2 message transparent transmission, perform the following step.

Command	Function
ZXR10(config-vfi)# **show vfi** <*vfi-name*>	This views a vfi detailed information including layer 2 message transparent transmission configuration policy.

This page is intentionally blank.

Chapter 5

MPLS OAM Configuration

Table of Contents

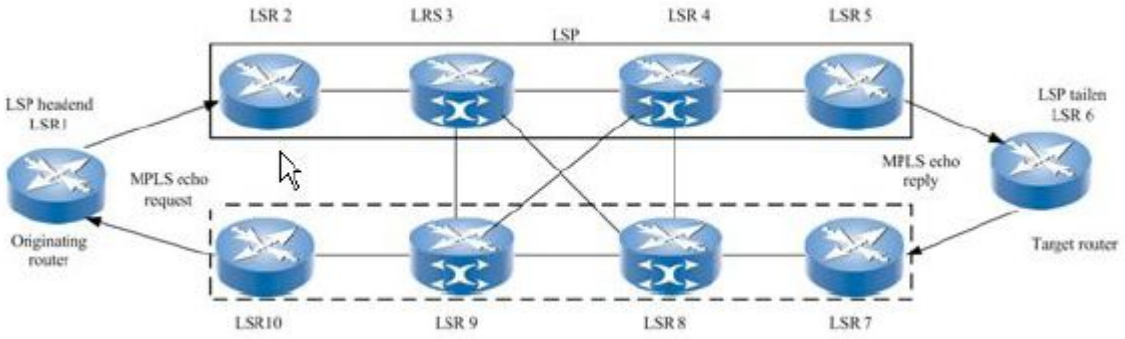
Function Introduction 65
Configure MPLS OAM 66

Function Introduction

MPLS TRACE/PING is used to detect MPLS LSP data plane fault. It is quite simple and practical. Router in internet can be used to report MPLS LSP data plane errors or provide information about unexpected events. This method can discover and isolate black hole routes, route loss, or other faults in short time.

MPLS TRACE/PING is an error report mechanism. It simulates ICMP request and response mechanism to finally locate network fault. Working principle of MPLS TRACE/PING is shown in Figure 33.

FIGURE 33 LSP TRACE WORKING PRINCIPLE



When initiating an LSP TRACE/PING request from LSR1, the destination is FEC (Forwarding Equivalence Class) in LSR6 and the route is shown in Figure 33. Firstly, LSR1 sends an MPLS echo request to LSR2. Since TTL field in outer tag is 1, after receiving this request, LSR2 finds that destination is not him. Therefore, it replies an MPLS echo reply packet with return to be 6. After receiving reply packet, LSR1 learns address of LSR2. LSR1 sends MPLS echo requests whose TTL values increment by 1. After receiving such

request, LSR5 takes out MPLS label and sends this echo request packet to LSR6 as a common IP packet. After receiving this echo request packet, LSR6 processes echo request and returns an MPLS echo reply packet with return value to be 3 to LSR1 along the reverse path. In this way, LSR1 can learn LSR addressees that LSR passes.

8900 series switches can use LSP TRACE/PING function to detect FEC (IPV4UDP and RSVP) of various types. LSP TRACE/PING request packet is a tagged UDP packet. As for this packet, the destination port is the well-known port 3503 and source port is specified by sender. Source address in IP layer is IP address of sender, destination address is 127.x.x.x/8, and TLV field contains downstream TLV to store label information for showing label information when 1SP TRACE/PING succeeds.

Configure MPLS OAM

Network node or link fault can be analyzed by executing command TRACE/PING.

1. Enabling/Disabling Show ICMP Debugging Information Function

Command	Function
ZXR10(config)#**debug mpls lspv** [**tlv**][**error**][**event**][**packet**]	This enables/disable show icmp debugging information function in global configuration mode. When the function is enabled, basic information of this echo packet can be shown when router sends or receives an MPLS echo packet. The function is disabled by default.

   Execute command **no debug mpls lspv** [**tlv**][**error**][**event**][**packet**] to disable show debugging information function.

2. Configuring trace mpls Command

Command	Function
ZXR10#**trace mpls** {**ipv4** *<destination-address>*[**destination** *<address-start><address-end><increment>*][**ttl** *<ttl>*]| **traffic-eng** *<tunnel-interface><tunnel-number>*[**ttl** *<ttl>*]}[*<source source-address>*][**reply mode** *<reply-mode>*]	This checks LDP of IPV4 and FEC of RSVP in privileged mode.

   Parameters:

Parameter	Description
< destination-address >	FEC type: LDP IPv4 prefix
traffic-eng	MPLS TE tunnel
destination	This is destination address, dotted decimal notation.
< source source-address >	This is source address.
< ttl >	This sets TTL value, ranging from 1 to 255. The default value is 30.
< reply-mode >	Response mode of reply packet: ipv4 or reply-alert

3. Configuring ping mpls Command

Command	Function
ZXR10#**ping mpls** {**ipv4** <*destination-address*>[**destination** <*address-start*><*address-end*><*increment*>][**ttl** <*ttl*>]| **pseudowire ipv4-address vc-id** <*vc-id*>[**destination** <*address-start*><*address-end*><*increment*>]| **traffic-eng** <*tunnel-interface*><*tunnel-number*>[**ttl** <*ttl*>]}[**source** <*source-address*>][**repeat** <*count*>][**timeout** <*seconds*>][{**size** <*datagram-size*>}][**reply mode** <*reply-mode*>][**interval** <*msec*>]	This checks LDP of IPV4 and FEC of RSVP in privileged mode.

Parameters:

Parameter	Description
< destination-address >	FEC type: LDP IPv4 prefix
traffic-eng	MPLS TE tunnel
pseudowire ipv4-address vc-id < vc-id >	This is AToM VC type.
< datagram-size >	This is the size of Ping mpls packet, ranging from 100 to 1500, 100 by default.
interval < msec >	This is timeout time (in seconds), ranging from 1 to 60.
reply mode reply-mode	Response mode of reply packet: ipv4 or reply-alert

Parameter	Description
< ttl >	This sets TTL value, ranging from 1 to 255.
destination	This is destination address, dotted decimal notation.

Chapter 6

MPLS TE HOT-STANDBY Configuration

Table of Contents

MPLS TE HOT-STANDBY Overview 69
Basic Configuration of MPLS TE HOT-STANGBY 69
Maintenance and Diagnosis of MPLS TE HOT-STANDBY 70
Configuration Example of MPLS TE HOT-STANDBY 70

MPLS TE HOT-STANDBY Overview

Hot-standby is the important function of MPLS traffic engineering. It strengthens the function of RSVP TE sub-system. Hot-standby is a kind of mechanism of redundancy link protection. It creates one protection route for every piece of protection lsp and provides protection ability for whole route.

Hot-standby can solve the failure problems about link, node, multi-link and multi-node in the complex network. Compared with traditional IGP handling method, hot-standby can switch the packet from invalid lsp to standby lsp, which reduces the loss of packets gradually. Compared with FRR which only can provide local protection, the end-to-end protection effect supported by hot-standby is better.

Basic Configuration of MPLS TE HOT-STANGBY

Hot-standby configuration is on tunnel interface of header node based on configuration of MPLS TE. The configuration command is as follows:

Command	Function
ZXR10(config-if)#**tunnel mpls traffic-eng hot-standby** [**protect** *<primary-path-option>* **secondary** *<secondary-path-option>*]	This configures MPLS TE HOT-STANGBY.

Parameter	Description
<primary-path-option>	path-option of protected master lsp
<secondary-path-option>	path-option of hot-standby lsp

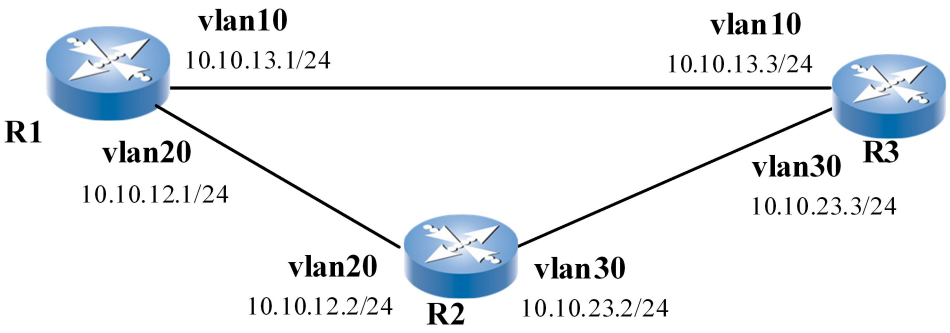
Maintenance and Diagnosis of MPLS TE HOT-STANDBY

ZXR10 8900 provides some commands to check working status of hot-standby which are as follows:

Command	Function
ZXR10(config-if)#**show mpls traffic-eng tunnels** [**tunnel id**]	This views the detailed information of hot-standby lsp.
ZXR10(config-if)#**debug ip rsvp hot-standby**	This is the diagnosis command of hot-standby.

Configuration Example of MPLS TE HOT-STANDBY

FIGURE 34 HOT-STANDBY CONFIGURATION EXAMPLE



The configuration task of the three routers in Figure 34 are shown as below:

Router	Loop-back	Task	Tunnel
R1	100.1.1.1	end node	Tunnel1, the destination address is 100.1.1.3
R2	100.1.1.2	middle node	
R3	100.1.1.3	end node	

Configuration of R1:

```
R1(config)#interface vlan10
R1(config-if)#ip address 10.10.13.1 255.255.255.0
R1(config-if)#ip rsvp bandwidth 30000 10000
R1(config-if)#mpls traffic-eng tunnels
R1(config-if)#exit
R1(config)#interface vlan20
R1(config-if)#ip address 10.10.12.1 255.255.255.0
R1(config-if)#ip rsvp bandwidth 30000 10000
R1(config-if)#mpls traffic-eng tunnels
R1(config-if)#exit
R1(config)#interface loopback1
R1(config-if)#ip address 100.1.1.1 255.255.255.255
R1(config-if)#exit
R1(config)#mpls traffic-eng tunnels
R1(config)#router ospf 1
R1(config-router)#mpls traffic-eng area 0
R1(config-router)#mpls traffic-eng router-id loopback1
R1(config-router)#network 100.1.1. 0.0.0.0 area 0
R1(config-router)#network 10.10.13.0 0.0.0.255 area 0
R1(config-router)#network 10.10.12.0 0.0.0.255 area 0
R1(config-router)#exit
```

Configuration of R2:

```
R2(config)#interface vlan20
R2(config-if)#ip address 10.10.12.2 255.255.255.0
R2(config-if)#mpls traffic-eng tunnels
R2(config-if)#ip rsvp bandwidth 30000 10000
R2(config-if)#exit
R2(config)#interface vlan30
R2(config-if)#ip address 10.10.23.2 255.255.255.0
R2(config-if)#mpls traffic-eng tunnels
R2(config-if)#mpls traffic-eng backup-path tunnel22
R2(config-if)#ip rsvp bandwidth 30000 10000
R2(config-if)#exit
R2(config)#interface loopback1
R2(config-if)#ip address 100.1.1.2 255.255.255.255
R2(config-if)#exit
R2(config)#mpls traffic-eng tunnels
R2(config)#router ospf 1
R2(config-router)#mpls traffic-eng router-id loopback1
R2(config-router)#mpls traffic-eng area 0
R2(config-router)#network 10.10.12.0 0.0.0.255 area 0
R2(config-router)#network 100.1.1.2 0.0.0.0 area 0
R2(config-router)#network 10.10.23.0 0.0.0.255 area 0
R2(config-router)#exit
```

Configuration of R3:

```
R3(config)#int vlan30
R3(config-if)#ip address 10.10.23.3 255.255.255.0
R3(config-if)#mpls traffic-eng tunnels
R3(config-if)#ip rsvp bandwidth 30000 10000
R3(config-if)#mpls traffic-eng tunnels
R3(config-if)#exit
R3(config)#interface vlan10
```

```
R3(config-if)#ip address 10.10.13.3 255.255.255.0
R3(config-if)#mpls traffic-eng tunnels
R3(config-if)#ip rsvp bandwidth 30000 10000
R3(config-if)#mpls traffic-eng tunnels
R3(config-if)#exit
R3(config)#interface loopback1
R3(config-if)#ip address 100.1.1.3 255.255.255.255
R3(config-if)#exit
R3(config)#mpls traffic-eng tunnels
R3(config)#router ospf 1
R3(config-router)#mpls traffic-eng router-id loopback1
R3(config-router)#mpls traffic-eng area 0
R3(config-router)#network 100.1.1.3 0.0.0.0 area 0
R3(config-router)#network 10.10.13.0 0.0.0.255 area 0
R3(config-router)#network 10.10.23.0 0.0.0.255 area 0
R3(config-router)#exit
```

Configure Tunnel on R1:

```
R1(config)#interface tunnel1
R1(config-if)#tunnel mode mpls traffic-eng
R1(config-if)#tunnel destination ipv4 100.1.1.3
R1(config-if)#tunnel mpls traffic-eng bandwidth 20000
R1(config-if)#tunnel mpls traffic-eng path-option 2 explicit-path
identifier 1
R1(config-if)#tunnel mpls traffic-eng path-option 4 explicit-path
identifier 2
R1(config-if)#tunnel mpls traffic-eng hot-standby
R1(config-if)#exit
R1(config)#ip explicit-path identifier 1 next-address 10.10.13.3 strict
R1(config)#ip explicit-path identifier 2 next-address 10.10.12.2 strict
R1(config)#ip explicit-path identifier 2 next-address 10.10.23.2 strict
R1(config)#ip explicit-path identifier 2 next-address 10.10.23.3 strict
```

Figures

Figure 1 IP Forwarding 4
Figure 2 ATM Forwarding 5
Figure 3 Position of MPLS..... 6
Figure 4 MPLS Working Principle 7
Figure 5 MPLS Label Structure..... 7
Figure 6 MPLS Label..... 8
Figure 7 MPLS Network Structure 9
Figure 8 Generating a Route Table 11
Figure 9 Generating a LIB.....11
Figure 10 Generating a LSP.....12
Figure 11 Penultimate Hop Popping 13
Figure 12 Downstream Unsolicited 14
Figure 13 Downstream on Demand 15
Figure 14 LDP Session Establishment 17
Figure 15 MPLS Basic Configuration Example.....21
Figure 16 Label Distribution Example 24
Figure 17 Structure of VPN-IPv4 Address 31
Figure 18 MPLS L3VPN Load Balance 34
Figure 19 Configuring OSPF in MPLS Domain 35
Figure 20 Configuring VRF on PE.....37
Figure 21 Configuring Static Route between CE and PE 38
Figure 22 Configuring RIP between CE and PE 40
Figure 23 Configuring OSPF between CE and PE.....41
Figure 24 Configuring EBGP between CE and PE.....43
Figure 25 Configuring MPBGP 44
Figure 26 MPLS L3VPN Configuration Example 45
Figure 27 Configuration Example of MPLS L3VPN Load Balance on Smartgroup Link 49
Figure 28 Network of VPLS 53
Figure 29 VPWS Configuration Example 57
Figure 30 VPLS Configuration Example 1.....58
Figure 31 VPLS Configuration Example 2.....60
Figure 32 VPLS Configuration Example 3.....61
Figure 33 LSP TRACE Working Principle.....65

Figure 34 hot-standby Configuration Example.....70

Glossary

ATM
- Asynchronous Transfer Mode

BGP
- Border Gateway Protocol

CE
- Customer Edge

DU
- Downstream Unsolicited

DoD
- Downstream-on-Demand

FEC
- Forwarding Equivalence Class

FIFO
- First In and First Out

FR
- Frame Relay

FRR
- Fast Reroute

IETF
- Internet Engineering Task Force

IPX
- Internetwork Packet Exchange protocol

IS-IS
- Intermediate System-to-Intermediate System

LDP
- Label Distribution Protocol

LER
- Label Edge Router

LIB
- Label Information Base

LSP
- Label Switched Path

LSR
- Label Switched Router

MPLS
- Multi Protocol Label Switching

NCP
- Network Control Protocol

OSPF
- Open Shortest Path First

PE
- Provider Edge

PPP
- Point to Point Protocol

QoS
- Quality of Service

RD
- Route Distinguisher

RFC
- Request For Comments

RSVP
- Resource ReSerVation Protocol

TCP
- Transfer Control Protocol

UDP
- User Datagram Protocol

VC
- Virtual Circuit

VCC
- Virtual Channel Connection

VCI
- Virtual Circuit Identifier

VPI
- Virtual Path Identifier

VPLS
- Virtual Private LAN Service

VPN
- Virtual Private Network

VPWS
- Virtual Private Wire Service

WAN
- Wide Area Network